Selecting Appropriate Game Factors in Educational Gamification: An Instrument for Investigating Undergraduate Students' Pleasurability in Learning

By

Gulinna A

Submitted to the graduate degree program in Educational Leadership and Policy Studies and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson Young-Jin Lee

_____

Ronald Aust

_____

Richard Branham

_____

Jennifer Ng

_____

John Poggio

Date Defended: 06.17.2016

The Dissertation Committee for Gulinna A

certifies that this is the approved version of the following dissertation:

Selecting Appropriate Game Factors in Educational Gamification: An Instrument for Investigating Undergraduate Students' Pleasurability in Learning

_____

Chairperson Young-Jin Lee

Date approved: 06.17.2016

Abstract

Gamification is defined as using game factors in non-game environments with the purpose of encouraging users to behave in a certain way. Educational gamification is to use game factors in learning environments. This innovative instructional/learning approach has positive effects on student engagement (Erenli, 2013; Hamari et al., 2014; Jensen, 2012; Nah et al., 2014). But most research used case studies with small sample size or less rigorous research methods. Since the results from those studies were contextual, they could not be generalized in diverse contexts. This mixed-method research aims to design an instrument for investigating undergraduates' pleasurability in educational contexts and then map them with game factors. This instrument could be used to inform instructors and designers of the most desirable game factors that can make students' learning experiences more pleasurable.

A 4-point Likert scale was distributed to 279 undergraduate students at the University of Kansas to explore student preferences for different types of pleasurable learning experiences. Four subscales were revealed from exploratory factor analysis (EFA) with the internal consistency of above .70 Cronbach's alpha scale reliability score. The correlations between undergraduates' pleasurability in educational contexts and the game factors in educational gamification were found based on the items that are included in each subscale and the definitions of the game factors.

Multivariate analysis of variance (MANOVA) was implemented to explore the relationships between participants' demographic, academic, and gaming/technology backgrounds on each subscale. *Gender*, *Ethnicity*, and *Frequency of Using Computers for Learning per Week* were the three predictors that resulted in statistically significant differences among college student senses of pleasurable learning experiences at the significance level of .05. This research

also conducted one-way analysis of variance (ANOVA) to compare Chinese international students with domestic students in the United States in terms of their pleasurability in educational contexts. Statistically significant differences were found between these two groups of participants.

The results of this research can be used for future research on designing alternative instructions for learners who come from different countries of origins. The instrument developed in this research will also help researchers design their experiments and evaluations on student engagement and pleasurability in gamified learning environments.

Acknowledgements

It has been such a long journey for me to get here. I cannot achieve this much without the helps and supports from my parents, my friends, and my dear professors. I want to dedicate this dissertation to my dad, another Dr. A in my family. He has always been my role model and has inspired me to study educational technology.

I also want to thank my mentor, Dr. Lee, who has educated and guided me for the past five years. Thanks to your patience and confidence in me, I could overcome all the obstacles on my way. Another very important mentor in my doctoral program is Dr. Meyen. Without your supports, I could not complete my doctoral program so smoothly. It has also been my great honor to know and to work with Dr. Poggio, Professor Richard, Dr. Aust, and Dr. Ng. Your knowledge and suggestions have helped me a lot in shaping and polishing my research.

My dear friends, thank you all for standing by me. Yang Zhao, Dan Chen, and Lu Wang, your ideas and insights have always inspired me. Last but not the least, thanks to all my other friends, your supports and companies have made my life much easier. I am very blessed to have you all.

Table of Contents

Introduction

Research Contexts

Educational environments have been significantly changed in the past two decades because of globalization and advances in technology. Online courses and mobile learning applications have emerged into the educational market. Technology-assisted teaching and learning approaches, such as Google classrooms, interactive whiteboard, and educational games, have also been implemented in educational environments. Besides the changes in educational environments, students' learning behaviors have also been reconstructed under this circumstance. Almost 97% of young people play games in their everyday lives (Erenli, 2013). They learn skills and methods through games in everyday lives, but they need to use another system of learning methods in academic domains (Erenli, 2013). Additionally, the population and the ethnical diversity of learners in higher education have also been tremendously expanded in recent years. The enrollment of international students in higher education in the United States has increased by 30.45% in the last decade (de Araujo, 2011).

All these new issues have added challenges to instructors and educators because they need to adjust their pedagogical strategies to meet the different learning needs of today's learners. How can instructors engage nowadays students in learning? Gamification, the use of game factors in non-game environments with the purpose of motivating users behave in a certain way (Deterding et al., 2011), is a possible solution to bridge the gap between learners' everyday learning habits and academic learning methods because it has the potential to motivate students to learn in educational environments.

Gamification is related to educational technology as its reward mechanics and dynamics are accomplished by computer software or website plug-ins. In information technology (IT)

research, Gartner estimates that more than half of the companies will implement gamification in managing innovation processes by the end of 2015 (Hamari, Koivisto & Sarsa, 2014; Xu, 2011). Higher education, as a critical stage that transfers the youth from educational environments to workplaces, should also be prepared to implement these gamified activities into learning and instructional contexts.

Statement of the Problem

Existing research on educational gamification has shown its positive effects on motivating students to learn in both K-12 and higher education (Erenli, 2013; Hamari et al., 2014; Jensen, 2012; Nah et al., 2014). But most research used case studies with small sample sizes or evaluated the effectiveness of those game factors by asking for user feedback *after* using the gamification (Nah et al., 2014) rather than investigating learners' needs and preferences before the design of gamification. Learner-oriented research is important because designing gamified instructions is also designing user interfaces. The technologies keep advancing because the users change the products when they interact with the products (Lindholm et al., 2003). A well-designed user interface should reflect users' needs. It is important to understand the clients holistically (Jordan, 2000). A systematic design and rigorous assessment of educational gamification is needed.

Purpose of the Study

This research developed an instrument to analyze university undergraduate students' pleasurability in educational contexts based on a game design framework named playful user experience (PLEX) by Arrasvuori and his colleagues (2011). The distinct feature of this research is that it found out undergraduate students' preferences for different types of pleasurable learning experiences and mapped those experiences with game factors in educational gamification. The

usefulness of this instrument is that it could tell the instructional designers and instructors about the game factors that their target students favor the most. Therefore, the design of gamified learning environments will be more effective.

According to Lindholm et al. (2003), "understanding the user is in effect understanding how that person performs the relevant tasks" (p. 90). Therefore, the tasks to be evaluated can be deconstructed into small pieces. This research focused on the initial step of designing gamified instructions, which was to select appropriate game factors in gamified activities. Jordan (2000) considers pleasurability as the highest level of product design, which refers to the pleasure with products: "the emotional, hedonic and practical benefits associated with products" (Jordan, 2000, p. 12). Understanding students' pleasurablity in educational gamification will maximize the design of comprehensive user interfaces.

Besides studying educational gamification through the lens of college students, this research also values the cultural perspective in studying educational gamification. In multicultural education, literature has shown that students from different cultural backgrounds have their own learning behaviors and habits. Instructors attempt to customize their instructions for all students instead of using one standardized teaching approach (Sleeter & Grant, 2009). "Culture" in this research was defined as students who came from different countries of origins. In the domain of business and interface design, cultural perspective has been highly valued. But in the domain of computer-mediated learning environments, researchers tend to think the learners are universal. This research aims to provide some insights for instructional designers and instructors with the idea of creating comprehensive user interfaces for students who come from different cultural backgrounds.

Research Questions

Two research questions have been investigated and answered throughout this research:

(1) *How can undergraduate students' pleasurability in educational contexts be reliably and robustly measured?*

   a. *What are the sub-constructs of pleasurability in educational contexts?*

   b. *How do different aspects of pleasurability in educational contexts correlate with game factors in educational gamification?*

(2) *What predictors significantly influence undergraduate students' pleasurability in educational contexts?*

   a. *In particular, whether, if there is any, undergraduate preferences for pleasurable learning experiences differentiate because of their countries of origins?*

Research Design

Both qualitative and quantitative research methods were implemented in the research design because of the exploratory orientation of this research. Qualitative methods, such as semi-structured interviews and focus groups, were conducted in the process of questionnaire item design with the purposes of understanding the experiences that undergraduate students regarded as pleasurable in educational environments. An expert panel, which was composed of experts in relevant instrument design and gamification research fields, reviewed and then finalized the description of each pleasurable learning experience.

An online Likert 4-point agree/disagree scale questionnaire (1: strongly disagree, 2: disagree, 3: agree, and 4: strongly agree) was distributed to the undergraduate students at the University of Kansas (KU) via Qualtrics. The participants were recruited by flyers and email invitations. A pilot study was administered with the purpose of checking the validity and

reliability of this instrument. 88 KU undergraduates participated in the pilot study. Exploratory factor analysis (EFA) was conducted to explore the underlying relationships among the Likert items and to finalize the items that should be included in a modified survey. Then the modified questionnaire was distributed to 279 KU undergraduate students. EFA was conducted to analyze the new list of Likert items used in that survey.

Multivariate analysis of variance (MANOVA) was conducted to analyze the measurement model. The dependent variables were the sub-constructs that were revealed from EFA and the independent variables were participants' demographic information, academic backgrounds, and technology/game backgrounds. The main effect of each independent variable on the dependent variables was discussed. As the Chinese international students were the second largest ethnical group that participated in the research, one-way analysis of variance (ANOVA) was used to compare the mean differences in each sub-construct between Chinese international undergraduates and the U.S. domestic undergraduates.

Significance of the Study

The objectives of this research are (1) to provide more empirical evidence for the existing literature on educational gamification, such as testifying whether gamer types, gender, and majors are the critical factors in the design of gamification; (2) to investigate whether Chinese international students have different preferences for game factors compared to the U.S. domestic students. This research is intended to provide some insights with the cultural awareness in studying international student preferences for the game factors in educational gamification. The results of this research can be used for future research on designing alternative instructions for learners who are from different cultural backgrounds. The instrument in this research will also

help future researchers build their instrument and experiments in studying student engagement and pleasurability in educational gamification.

Literature Review

Economic globalization and technology advancement have brought significant changes to higher educational environments. The enrollment of international students has been increased and the strategic development of higher educational institutions has also been influenced by internationalization. The target subjects of this research were university undergraduates because graduate and undergraduate students have different academic motivations. Graduate students tend to be more self-motivated in learning while the undergraduates still explore their interests and academic goals in learning.

This section will start with the definitions of gamification and game factors in gamification, which are the critical concepts in educational gamification. Some newly developed educational theories and philosophies will also be introduced with the purpose of understanding current university undergraduate students' learning behaviors and engagement issues in the United States. Current empirical evidences on the effectiveness of using gamification as a teaching/learning approach to engage students will be discussed in this section as well. Then this section will discuss the importance of understanding pleasurability in designing educational gamification and the design frameworks that were used in this research.

What is Gamification?

Gamification is defined as the use of game factors in non-game environments with the purpose of encouraging users to behave in a certain way (Deterding et al., 2011). Some people mix the definition of gamification with game-based learning. Game-based learning operates based on simulations and games that have educational purposes. Gamification is different from serious games or simulations in the fact that gamification does not necessarily involve playing games. Gamification only uses the fun and pleasurable game mechanics that are found in serious

games. The majority of people nowadays have all experienced gamification to certain extent, such as using social network sites (SNSs) and attending massive online open courses (MOOCs). For example, in Khan Academy, users can track their progress and preview future tasks from a visualized progress tree. LinkedIn can be also considered as an example of gamification because it works as an SNS with the purpose of encouraging users to build wider connections with peers, colleagues, and the scholars in the similar fields of studies by providing game factors such as profile, progress bar, and rewarding mechanics, e.g., endorsements for users' claimed skills.

Game Factors in Gamification

Most gamified activities include three basic parts: "goal-focused activity, reward mechanisms, and progress tracking" (Glover, 2013, p. 2000). The operation system of gamification works similarly to instructions in education – learning and teaching objectives, assessments, and syllabus. These three parts of gamification makes it possible for adoption as a teaching/learning approach in educational environments. These three parts of gamified activities can be broke down to the design of seven game factors (Glover, 2013; Hamari et al., 2014; Kapp, 2012):

- *Storytelling:* It works as the rules of the gamified activities. A good gamification should have clear and simple storyboard to direct the users to achieve the goals. This element functions like the guidelines and directions of an instructional activity in class.
- *Levels:* A gamified activity usually consists of different levels for users/learners to advance. In each level, users/learners will face different challenges. These levels and challenges can be viewed as learning objectives for learners.
- *Points:* It belongs to the progress tracking part because according to the rules (storytelling), users/learners can gain points when they complete the gamified activities.

- *Leaderboard:* This element is under reward mechanism, which shows the users/learners who are leading in the gamified activities. This element is very controversial when the gamification is used in educational contexts because some empirical evidence shows that leaderboard only suits for those people who are aggressive and hardcore players (Harmari et al., 2014).
- *Badges:* This element works as the visualization of rewards that users/learners have achieved when they accomplish certain achievements. This element works as the extrinsic motivation for users/learners (Kapp, 2012).
- *Feedback*: A well-designed gamification interface should provide users/learners with timely feedback in order to help them to stay on the right track.
- *Progress:* A progress-tracking bar should appear in the user/learner profile in order to remind users/learners how many tasks are left and how many accomplishments they have achieved.

The Development of Gamification

The past two decades have experienced unprecedented technology advancement, including in the field of education. Educational technology, such as online learning and mobile learning, assists instructors in higher education to better meet learners' diverse needs. More and more instructors use online collaboration software to share learning and supplementary materials with students, such as Wikis and Learning management systems (LMSs), e.g., Blackboard and Canvas. Under this big environment, gaming theory, which is the study of strategic decision-making, has also been adopted to educational contexts in order to engage learners (Landers & Callan, 2011).

Although computer games and simulations have been implemented into learning and instructional environments for a fairly long time, educational gamification is a relatively new concept, which refers to the application of game factors in non-game activities in education. This innovative approach is connected to educational technology as the game factors of gamification are accomplished through computer software and/or website plug-ins. It is critical to develop a system to better understand the relationships between the game factors and desired learning outcomes (Pedreira et al., 2015).

Besides the gaming theory, there are also some learning theories and pedagogical theories behind educational gamification. For example, the rewarding mechanism in gamification is aligned with behaviorist theory, which is to directly change users' behaviors through motivation and rewards (Schultz, 2006). The feedback and storytelling elements in gamification are aligned with constructivist theory, which is learner-centered and aims to encourage collaborations (Bonk & Cunningham, 1998).

Gamification in E-Learning in Higher Education

The definition of e-learning is the use of information and communication technologies in improving students learning performances (Higher Education Funding Council of England, 2005). User satisfaction is one of the most important factors in assessing the success of system implementation (DeLone & McLean, 1992). According to Urh et al. (2015), the research on e-learning in higher education focuses on student characteristics, such as their conceptions of learning; course context, such as teaching methods; learning context, such as student perceptions of the quality of teaching and the quantity of work; student approaches to learning, such as what they do and why they approach learning in particular ways; and the quality of student learning outcomes.

Besides the characteristics of research foci in e-learning in higher education, the structures of higher education itself have also been changed significantly in recent years given that more and more inter- and multi-disciplinary programs have been emerged, as well as the increasing number of participations in international mobility programs bring about specific learning situations (Biro, 2013; Urh et al., 2015). These characteristics of new e-learning in higher education can be handled more effectively based on gamification theory.

Well-designed e-learning that uses gamification can enhance learner satisfaction and engagement, and learning effectiveness and efficiency. The right combination of e-learning and gamfication can bring the state of flow to learners (Urh et al., 2015). Csikszentmihalyi (1990) defines *flow* as an optimal experience characterized as a state of being fully focused and engaged in an activity. According to McGonigal (2011), this feeling could be triggered by four game factors that are shared in well-designed games: feedback, goals, rules, and voluntary participation. The state of flow is important for the learners in conducting challenging activities as they would concentrate on the projects and manifest their skills optimally.

Internationalization in Higher Education

The subjects targeted in this research were the undergraduates in higher education in the United States. Advances in technology, such as MOOCs, also bring about demographic changes in student population in higher education. As a byproduct of globalization, higher educational internationalization has provided more opportunities and competitions for global learners and the domestic institutions in the United States. More and more young people are eager to study abroad in order to achieve higher human, social, and cultural capitals. The rapid increase of international students has also brought enormous revenues to the United States. By 2004, international students have contributed $12 billion annually to the United States although they

only constitute 4% of America's total college students (Zhai, 2004). On the other hand, the institutions in the United States are enthusiastic in building themselves as world-class universities (WCU) with the purpose of gaining worldwide prestige. One of the standards of being a WCU is the recruitment of international students (Glover, 2013; Lee, 2010). These push-and-pull factors for higher educational internationalization have accelerated the institutions to adapt their curricula and policies more friendly to international students.

University Student Engagement in the 21st Century

In order to help international students adapt to new learning environments while helping domestic students in the United States understand their international classmates, many research focuses on studying student engagement. The most popular definition of student engagement is from Kuh (2009), which is derived from Astin's (1985) theory of student involvement. It "represents the time and efforts students devote to activities that are empirically linked to desired outcomes of college and what institutions do to induce students to participate in these activities" (Kuh, 2009, p. 683).

In the process of international students' adjustment to the academic learning environments in the United States, they have encountered challenges in different degrees, such as language barriers, less social supports, and less engagement in active learning (Andrade, 2006; Krause, 2005; Lee, 2010; Xiang & Shen, 2009). Students experience learning through being involved (Krause, 2005). Studying student engagement is critical because it has positive relationship with students' academic achievements (Horstmanshof & Zimitat, 2007; Krause & Coates, 2008; Kuh, 2009). Krause (2005) concludes that measuring student engagement should evaluate from five dimensions: class contents, institutes, peers, online learning, and staff.

Online learning is a new evaluation dimension compared to the studies on student engagement in late 20th century. This dimension indicates that learner behaviors have been changed because their daily lives have been significantly changed by technology advancement. Today's learners on average spend 4.2 hours per week on the Internet for study and research purposes (Krause, 2005). Erenli (2013) argues that the learners learn skills and methods through games in everyday lives but they need to use another system of learning methods in academic domains. Gamification seems to be one possible approach to narrow the gap between today's learners' everyday learning experiences and academic learning methods.

Kuh (2009) suggests that a good undergraduate education experiences should include: "(a) student-faculty contact, (b) active learning, (c) prompt feedback, (d) time on task, (e) high expectations, (f) respect for diverse learning styles, and (g) cooperation among students" (p. 684). These standards can be accomplished by the game factors in gamification by using different levels of challenges to engage students while providing rewards and competitions for students to learn from each other.

From faculty perspective, they also need to make cognitive models visible to engage students (Chen, Lattuca, & Hamilton, 2008), which suggests that the assignments and in-class tasks need to be customized to students' cognitive capabilities. With the pursuit of achieving the state of flow in learning, which is a statement that students are fully motivated by the learning tasks, faculty members need to find the balance between students' academic achievements and the enjoyments of the tasks (Chen et al., 2008). By doing this, the learning tasks will not overload students' cognitive systems. Meanwhile, instructors also need to use collaboration systems to engage the interactions between the U.S. domestic students and international students. Gamification can help faculty members to achieve these teaching objectives easily as the design

of gamification is aligned with flow theory and one of the objectives of gamification is collaboration (Glover, 2013; Hamari et al., 2014; Nah et al., 2014; Xu, 2011).

Cultural Awareness in Business & Design vs. Educational Contexts

From the theories in multicultural education, students from different cultural backgrounds have their own learning behaviors and habits. Pedagogical approaches can be customized according to the differences among the student population in classrooms. However, these learning habits are mainly observed from face-to-face instructional contexts. In technology-mediated instructional environments, such as online learning environments, the learning preferences of students who come from diverse backgrounds are somewhat overlooked. In contrast, in the domain of business and design, cultural perspectives have been highly valued.

Hofstede and Hofstede (2005) define cultures "consist as the unwritten rule of the social game" (p.3). They provided a model for analyzing organizations for cultural differences and then extended it to an examination of culture from studying differences among countries around the world. According to Hoft (1995), this model focuses on the patterns of thinking, feeling, and acting that form a culture's mental programming. Researchers and designers have used this model to design user interfaces and other interactive products (Jordan, 2000). Barber and Badre (1998) define people's preferences on visual stimuli as cultural marks, which indicate that people from different cultural backgrounds have their specific visual preferences. They find that people from different countries have different preferences when they browse websites, such as colors, pictorial images, and text orientations. It is important for instructional designers to find user interface design layouts that work for most people. Barber and Badre (1998) conclude that both cultural markers and genres influence the website design. Culturability, being defined as cultural preferences and biases (Barber & Badre, 1998), should be one of the design principles in website

design. An international user interface should take cultural markers into consideration when the interface is designed for particular international users.

In the fields of business and design, they define cultures as specific attributes for the elements they would measure, but in education, the definition of culture is vague. In international technical communication, they use an approach named international-user analysis to understand the different needs of users who come from different countries (Hoft, 1995). In technology-mediated learning environments, the features that are used to accommodate learners' learning preferences are not well developed. This research recruited undergraduates from different countries of origins to compare whether they have different preferences for pleasurability in learning.

Empirical Evidences on the Effectiveness of Educational Gamification

Most research on educational gamification studies the effectiveness of each game factor in its influence on learner outcomes (Nah et al., 2014). The majority of these empirical studies on educational gamification show the positive relationship between gamified instructions and learner outcomes, particularly in improving students' class participation and motivation (Erenli, 2013; Hamari et al., 2014; Jensen, 2012; Nah et al., 2014). However, there is little research on studying game factors in educational gamification from cultural perspective in higher education. Both quantitative and qualitative methods have been used for data analysis in the literature, but some research has small sample sizes and less rigorous methodologies, such as lack of control groups and short experiment timeframes (Nah et al., 2014). Additionally, little research has used advanced statistical analysis approaches such as structural equation modeling (SEM).

Gamification has positive effects on learner outcomes only when the design of gamified instructions is customized to the users, e.g. majors, genders, and prior knowledge of learning

materials (Nah et al, 2014). Most research studies science, technology, engineering, and mathematics (STEM) majors whereas little studies the fields of social science. Dominguez et al. (2013) suggest that educational gamification fosters high-order thinking, e.g. problem-solving, rather than factual knowledge. But if gamified instructions can only benefit instructors in STEM majors, how can students in social science experience engaging learning contents through gamified instructions? What activities and learning supports do students in social science need?

These questions need in-depth observation and research. If gamification can suit current learners' needs and learning habits, it is the instructional designers' and educational practitioners' responsibility to find a way to design effective gamified instructions for students with diverse needs and who come from diverse backgrounds. A mature and intuitive design framework for educational gamification needs to be built in order to investigate learners' diverse needs for pleasurability in educational gamification.

Pleasurability

Jordan (2000) concludes the effectiveness of product design into three categories: functionality, usability, and pleasure.

Level 1: Functionality – refers to the necessary functions of a product to perform tasks

Level 2: Usability – refers to the how easy it is to use the product

Level 3: Pleasure – refers to the satisfaction with the use of the product that users can emotionally relate to their real lives

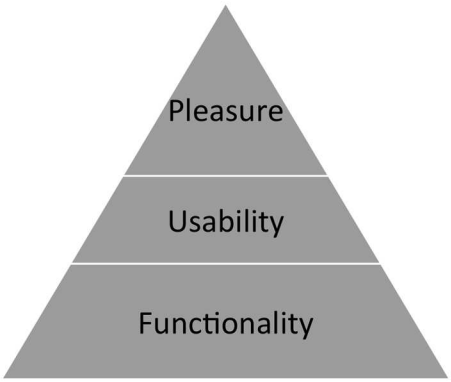

Figure 1. Modified chart of "A hierarchy of consumer needs" from Jordan, 2000, p. 6.

Pleasurability of a product can be further categorized into four types: physio, socio, psycho, and ideo (Tiger, 1992). Physio-pleasure refers to the pleasures derived from the sensory organs, such as a hug and body touch; socio-pleasure refers to the pleasures from relationships with others, such as conversations with other people and others' company; psycho-pleasure refers to the satisfaction that "in the act and in using the skill, energy, and resources to complete it" (Tiger, 1992, p. 56), such as the joy of completing some major tasks; and ideo-pleasure refers to the pleasures that in people's brains. One kind of ideo-pleasure is "receiving from experiencing or creating theoretical entities" (Tiger, 1992, p. 59), such as playing crossword puzzles or watching movies; the other kind of ideo-pleasure is found from the general impact of natural circumstance, such as the need of growing plants in the workplaces or keeping pets at home. In the case of designing game factors of educational gamification, these four aspects of pleasure can be specified to the chart below (See Table 1):

Table 1

Four-pleasure Analysis for Pleasurability in Ggame Factors

Physio	*Socio*	*Psycho*	*Ideo*
• Points • Progress Tracking	• Collaboration • Feedback	• Challenges • Levels	• Badge • Leaderboard • Storytelling

In learning undergraduate students' pleasurability in learning, this research used qualitative methods to explore the learning experiences that undergraduate students consider as pleasurable. Each pleasurable learning experience was presented in the format of a Likert 4-point agree/disagree item. An expert panel, which was composed of experts in instrument design and gamification, reviewed the clarity and conciseness of the description of each pleasurable learning experience. All qualitative research activities were conducted after granting the approval from Human Subject Committee at KU Lawrence campus (HSCL).

Design Frameworks

Several research and design frameworks will be introduced in this section. These frameworks were the guidelines that were used in conducting this research. As this research was to create an instrument for understanding undergraduate students' pleasurability in learning environments and then to match the pleasurable learning experiences with game factors in educational gamification, interdisciplinary research approaches were implemented, e.g., the research methods in interaction design and the design framework in game development.

Activity Theory

Educational gamification is presented to learners by technology-mediated learning modules such as via online courses or educational software. It is important to understand how learners perceive pleasurability in gamification through their previous learning experiences. Activity theory (See Figure 2) that was raised by the Russian psychologist Lev Vygotsky was the critical design theory to guide the empirical research activities in this research. The strength of activity theory is that it is a holistic design framework that involves multiple dimensions in conducting and evaluating human engagement with the world (Gay & Hembrooke, 2004). As

learning is a social behavior, it is necessary to analyze the learners' preferences for different types of pleasurable learning experiences from a comprehensive perspective.

The *Subject* is the individual or group of individuals in the activity. The *tool* includes the artifacts that can act as resources for the subject in the activity. The *object* is the goal of the activity. The *rules* are any regulation that can influence how the activity takes place. The *community* is the social group that the subject belongs to while engaged in an activity. The *division of labor* refers to how the tasks are shared among the community (Yamagata-Lynch, 2010).

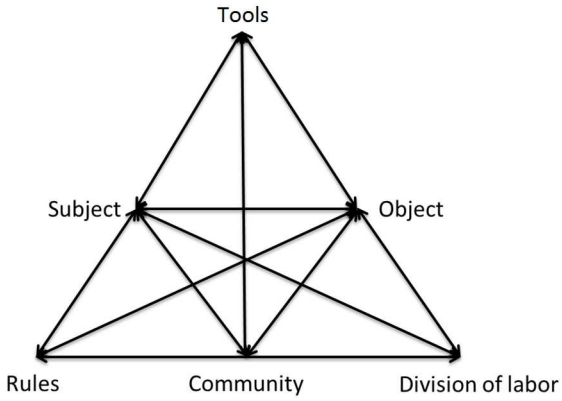

Figure 2. Originally Figure 1.1 Engerstrom's analysis of activity and mediating relationships, Gay & Hembrooke, 2004, p. 3)

Based on activity theory, activity systems analysis method has been used to evaluate complicated learning environments, such as how students interact with technology in classes (Yamagata-Lynch, 2010). Activity systems analysis method is important for qualitative research because it is designed to enhance understanding of human activity situated in a complex setting and is graphically represented by a series of triangles diagrams. By using these triangle diagrams, researchers and audience can understand the relationships between each element and the influence of each element to the other elements.

In order to interact with learners and to understand their pleasurable experiences in educational contexts, this research implemented activity systems analysis method in conducting focus groups when designing questionnaire items. Figure 2 is adapted to the model below (Figure 3). The usefulness of activity systems analysis method is that it shows the logic of conducting qualitative research in understanding a situated environment (Yamagata-Lynch, 2010). This method helps researchers sort and interpret qualitative data from the interviews and focus groups.

In this research, the activity that needs to be investigated is how students learn in educational contexts. *Undergraduate students* are the subject in this activity. The *class interactions* are the tool to understand this activity. Having students experience *pleasurable learning experiences* is the goal of this activity. The *learning objectives* in each different courses are the regulations that affect this activity. The community of this activity is the *online and/or face-to-face learning environments*. The division of labor in this activity is the *relationships between the students and the instructor*.

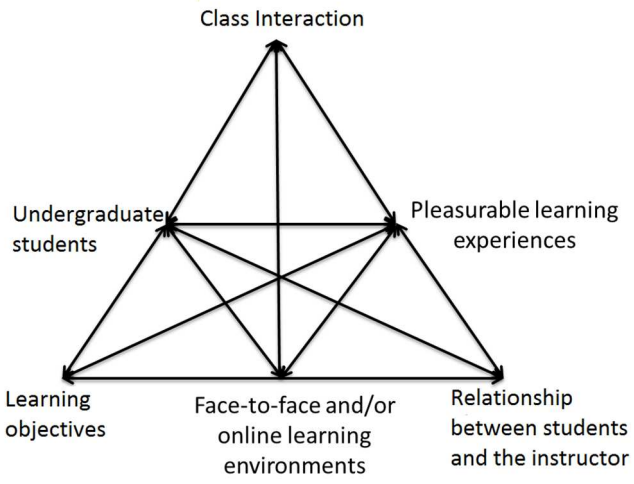

Figure 3. Activity system analysis method in evaluating pleasurability in educational gamification (adapted from Yamagata-Lynch, 2010)

Playful Experiences (PLEX)

A game design framework named playful experiences (PLEX), which has been used to guide game design (Arrasvuori et al., 2011; Kim, 2013; Korhonen et al., 2010), was also used in the process of questionnaire item design. PLEX is a design framework built upon the collection of several previous game design frameworks, which classifies the fun and playful factors into 22 categories by Nokia research team (See Table 2). The research team has extended this design framework to PLEX cards in order to put this framework into practical use. Two design techniques have been used with PLEX cards, *PLEX Brainstorming* and *PLEX Scenario* (Arrasvuori et al., 2011).

Table 2

PLEX framework and its 22 categories from Arrasvuori et al. (2011)

Experience	**Description**
Captivation	Forgetting one's surroundings
Challenge	Testing abilities in a demanding task
Competition	Contest with oneself or an opponent
Completion	Finishing a major task, closure
Control	Dominating, commanding, regulating
Cruelty	Causing mental or physical pain
Discovery	Finding something new or unknown
Eroticism	A sexually arousing experience
Exploration	Investigating an object or situation
Expression	Manifesting oneself creatively
Fantasy	An imagined experience
Fellowship	Friendship, communality or intimacy
Humor	Fun, joy, amusement, jokes, gags
Nurture	Taking care of oneself or others
Relaxation	Relief from bodily or mental work
Sensation	Excitement by stimulating senses
Simulation	An imitation of everyday life
Submission	Being part of a larger structure
Subversion	Breaking social rules and norms
Suffering	Experience of loss, frustration, anger
Sympathy	Sharing emotional feelings

Thrill	Excitement derived from risk, danger

Each PLEX card has a visual and textual content on one side, and the other side is designed with logo and patterns as the back of a common poker card. In a *PLEX Brainstorming* activity, three cards are drawn randomly from the deck to create an idea by the designer. Then the participants take turns in drawing a card from the deck and state how this experience category manifests in the idea. In a *PLEX Scenario* activity, three cards with background templates and actions, e.g., "beginning – who are the people in the story? How does this category launch the story?"; "continuation – How does this category cause the story to continue in a new direction?"; and "the end – How does this category bring the story to a close?", are provided to the participants. Participants create a scenario using the three cards selected from a set of seven or more available cards. The scenario is triggered by the actions related to the three cards with background templates. Arrasvuori and his colleagues (2011) evaluated the effectiveness of PLEX cards both quantitatively and qualitatively with design students in a five-month course started October 2009 at the Utrecht School of the Arts in the Netherlands.

Kim (2013) used this framework to design a gamified instructional module for 30 undergraduate students at engineering school of K university in South Korea. He conducted a Likert-scale online survey based on the categories in PLEX to analyze the students' needs on fun and pleasure. The top five preferences by these 30 engineering students were challenge, exploration, relaxation, completion, and discovery, while competition, eroticism, suffering, sadism, and control were relatively less preferred by this group of students.

These empirical evidences have shown the effectiveness of PLEX and the possibility of using such a design framework in investigating undergraduate students' pleasurability in learning. These 22 categories of fun and playful factors in PLEX were used as the guidelines and

references for conducting qualitative inquiries with research participants, as well as in generating the descriptions of pleasurable learning experiences. PLEX cards were used in the focus groups for brainstorming. Participants flipped the cards to discuss when and how they experienced that particular pleasurable feeling in learning environments.

However, as PLEX framework is originally used for general game design (Arrasvuori et al., 2011), some of the categories in PLEX are not suitable in educational contexts, for example, cruelty, eroticism, subversion, suffering, and thrill. On the other hand, some pleasurable learning experiences in formal educational contexts are not included in PLEX framework. Two new categories of pleasurable learning experiences were defined by the participants in this research: *Visualization* and *Interaction*.

Maxwell (2013) indicates that qualitative methods are very useful when the researchers intend to understand the real problems and situations. Interviews and focus groups were implemented in order to find out the pleasurable learning experiences in educational contexts. The mapping of student pleasurable experiences in learning contexts and game factors have been collected and designed as the chart below (Figure 4):

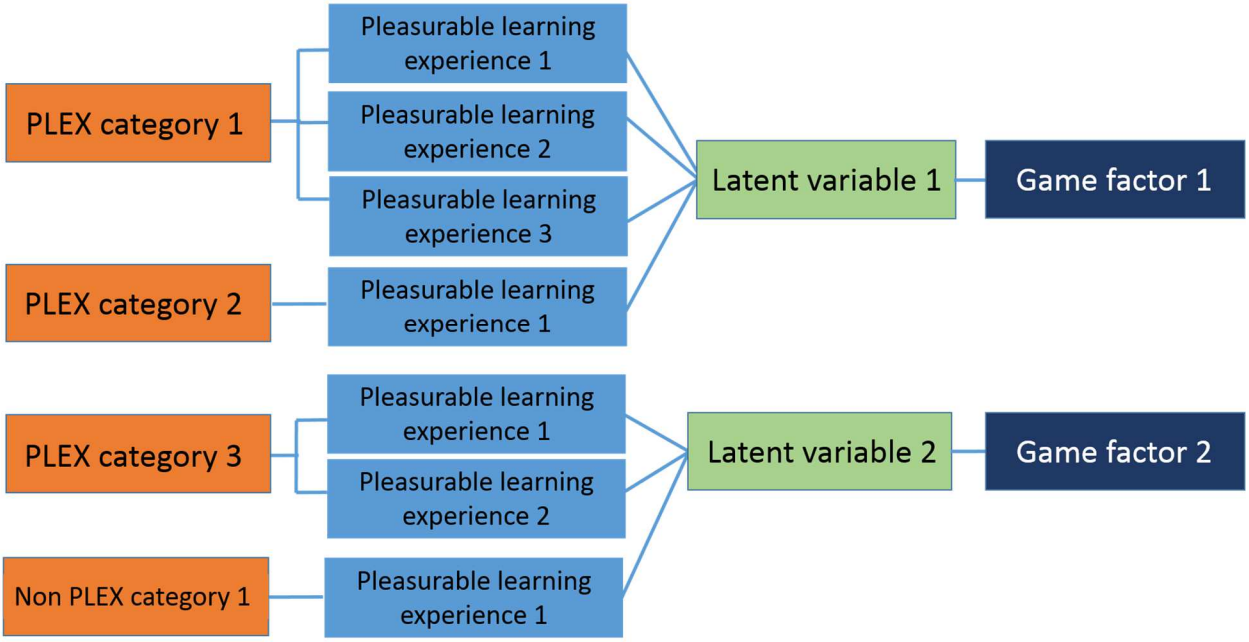

Figure 4. Mapping chart of pleasurable learning experiences and game factors

Research Questions

Based on existing literature, the factors that can influence the design of gamification in educational contexts are not clear. A learner-oriented research model need to be conducted in order to design comprehensive gamified instructions for learners. Two research questions have guided the design of this research:

(3) *How can undergraduate students' pleasurability in educational contexts be reliably and robustly measured?*

   c. *What are the sub-constructs of pleasurability in educational contexts?*

   d. *How do different aspects of pleasurability in educational contexts correlate with game factors in educational gamification?*

(4) *What predictors significantly influence undergraduate students' pleasurability in educational contexts?*

b. *In particular, whether, if there is any, undergraduate preferences for pleasurable learning experiences differentiate because of their countries of origins?*

Methodology

This research implemented both qualitative and quantitative research methods to understanding undergraduate students' needs for the selections of game factors in educational gamification. In this section, critical design concepts and frameworks will be explained. The methods of inquires that were used in this research and the guidelines of using those measurements will also be illustrated. Multivariate analysis techniques including EFA and MANOVA will be described in the statistical analysis part.

The main purpose of this research was to create an instrument – Likert 4-point agree/disagree scale questionnaire – to investigate university undergraduate students' pleasurability in learning and then map these experiences with the deign of game factors in educational gamification. The development of this scale consisted of two parts:

(1) *Likert item design.* This part included qualitative research methods, e.g., interviews and focus groups. The purpose of using qualitative research methods in this step was to explore undergraduate students' real pleasurable learning experiences in educational contexts. After conducting these two qualitative inquires and comparing with PLEX categories, a pool of Likert items has been generated. Each Likert item was a description of a pleasurable learning experience. These survey items were as specific as possible in order to reflect the scale's purpose (DeVellis, 2003).

When writing the Likert items, three or four times more than the number of items that were anticipated to be included in the final instrument were generated. DeVellis (2003) indicates that it is better to be over-inclusive when developing a scale. Then standards of item selections, such as the length and clarity of wordings, were applied. According to DeVellis (2003), scale developers should avoid exceptionally lengthy items and should consider about the reading

difficulty levels when writing items. Most instruments use the reading difficulty level at a fifth- to seventh-grade reading level. A typical item in this scale contains 9 words and 15 syllables. An expert review was conducted in order to check the clarity and conciseness of each Likert item – the description of a pleasurable experience in educational environments.

(2) *The distribution and the analysis of the online questionnaire*. An online 4-point Likert scale questionnaire with the items generated in the first step was distributed to voluntary and anonymous participants via on-campus flyers and social media tools. A pilot test with 88 participants was conduct in order to finalize the survey items and to understand the underlying relationships among the survey items. Then a modified survey was distributed to 279 undergraduates at KU.

Several data analysis techniques were used to analyze the collected data. Descriptive statistics, such as the frequencies of each category in the independent variables, mean score and the standard deviation for each subscale in the survey were computed. EFA was conducted to answer Research Question 1, which was to explore the relationships between pleasurable learning experiences and game factors in educational gamification. MANOVA was also conducted in order to answer Research Question 2, which was to analyze the relationship between the independent variables and the dependent variables. In particular, one-way ANOVA was conducted between Chinese international undergraduates and the U.S. domestic undergraduates with the purpose of further analyzing the differences in their preferences for the pleasurable learning experiences. This comparison was used to answer Research Question 2a.

Instrumentations

Both qualitative and quantitative research methods were used in this research. The introduction of each method follows the two steps of the design of this research (1) designing

questionnaire items, and (2) distributing the questionnaire. Qualitative research activities such as semi-structured interviews, and focus groups were conducted in Step (1) to design questionnaire items. Then a 4-point Likert scale questionnaire was distributed to KU undergraduates in Step (2) to investigate students' preferences for different types of pleasurable learning experiences.

Semi-structured interviews

The purpose of interviews was to explore the pleasurable learning experiences that students from different majors shared in both face-to-face and online educational contexts. Seven current undergraduates and graduate students at KU were interviewed. They came from different majors, e.g., business, design, education, engineering, journalism, mathematics, and psychology. These interviewees were recruited by email and personal contacts. The target undergraduate students recruited in the interview were those who have studied at KU for at least three semesters. Graduate students were those who had teaching experiences with undergraduates at KU (GTAs). The reason for including GTAs in the interviews was to understand the real classroom interactions in a comprehensive way. Both of the undergraduates and the GTAs were the stakeholders in the learning environments. The interview protocols for these two groups of interviewees were provided before conducting the interviews (See Appendix A).

Each interviewee signed a consent form before participating in the interview. Each interview lasted approximately one hour. The format of interview questions was open-ended rather than leading questions. All interviews were audio recorded and transcribed into written scripts. The transcript of the interview was shared with the respective interviewee for member check. All interview transcripts were stored in a secured computer with password for two years after taking the interviews.

Focus groups

Two groups of undergraduates and GTAs at KU were invited to discuss the "Likes and Dislikes" among their current educational experiences in face-to-face courses and online courses. The reason for including GTAs in the focus groups was to understand the situation from a comprehensive way. Each group contained 3-4 participants. One focus group only consisted of Chinese students and the other focus group was U.S. domestic students. The goal of conducting focus groups was to explore whether there were other shared pleasurable learning experiences besides the items generated from the interviews. Interactions between participants should be encouraged and thus, participants may have more opinions to share (Jordan, 2000). Focus groups also aimed to discover if there were different preferences between Chinese and U.S. participants.

The researcher debriefed the purpose of focus groups and the participants signed the consent forms before they began the discussions. Each participant listed their pleasurable learning experiences in face-to-face instructional environments and online courses separately (See Appendix B). Frequently mentioned pleasurable experiences in the interviews and PLEX brainstorming cards were presented to the participants in the focus groups for brainstorming. The participants could refer to those PLEX categories or create new categories and then wrote down their pleasurable learning experiences on Post-It note stickers. The participants shared their pleasurable learning experiences with the other participants in the groups. The results of their shared experiences were generated as new questionnaire items in the questionnaire. The discussions being occurred in focus groups were audio recorded.

Questionnaire

The format of this questionnaire was an online Likert 4-point agree/disagree scale questionnaire via Qualtrics – 1: strongly disagree, 2: disagree, 3: agree, and 4: strongly agree – with 22 descriptions of pleasurable learning experiences in educational contexts in the finalized

survey. For example, “I like collaborating with my classmates”; “I like reviewing contents”; “I like building connections with instructors”. Participants chose from a 4-point agree/disagree scale to rate their agreements with each statement. Besides responding to the Likert items, participants also provided their demographic information, academic background, and gaming experiences in the survey. An information statement was presented to the participants before they took the online questionnaire via Qualtrics with the purpose of introducing the survey and obtaining participants’ approvals. No personal identifiable data were collected in this research.

Descriptions of Variables

The dependent variable and independent variables are discussed in this session. These variables were used in the design of the online Likert scale questionnaire and data analysis. There are two parts in the survey. The first part is about participants’ personal backgrounds, which are the three sets of independent variables. The second part of the survey is the list of Likert items, which is used as the dependent variables in data analysis.

Dependent variable

For Research Question 1, the dependent variables were the list of Likert items. In the pilot test, there were 34 Likert items; in the modified survey, there were 30 items. Based on the EFA results, the finalized survey consisted of 22 Likert items and could be categorized into four subscales. For Research Question 2, the dependent variables were the four factors that were revealed from EFA.

Independent variables

Three sets of independent variables were collected for the online survey. The format of these questions was multiple-choices questions. The first set was questions about students’

personal information. Their age, ethnicity, and gender were collected. The detailed criteria and guidelines are listed as below:

- *Age*: As the participants in this research were university undergraduates, the common age range should be from 18 to 22. This research was anticipated to distribute to adult learners who were over 18 years old. Thus, the first category of age choice was 18-19; the second category was 20-21; the third category was 22-23; and the fourth was 24+.
- *Ethnicity*: The categories of this independent variable were based on 2010 US Census. Two questions were generated for this variable. One was asking students' Hispanic, Latino, or Spanish origins. And the other gave students' choices to identify their specific ethnical backgrounds. These categories were: 1) White; 2) Black, African Am., or Negro; 3) American Indian or Alaska Native; 4) Asian Indian; 5) Japanese; 6) Native Hawaiian; 7) Chinese; 8) Korean; 9) Guamanian or Chamorro; 10) Filipino; 11) Vietnamese; 12) Samoan; 13) Other Asian _____; 14) Other Pacific Islander _____; 15) Some other race _____. The difference between 2010 Census and past Census question for race and ethnicity is that 2010's is more specific and focuses more on people's self-identification.
- *Gender*: As shown in Kim's (2013) research and some other research on gaming (Kapp, 2012, Koivisto & Hamari, 2014), gender is an important factor that affects users' preferences for fun and pleasure in gamification. The categories under this independent variable are Male and Female.

The second set of independent variables was questions about students' academic backgrounds. According to previous research, undergraduates' current years in their colleges and majors have influences on students' academic engagement and their preferences for gamified

instructions. The criteria and guidelines for the categories under each independent variable are listed as below:

- *Current studying year*: Some research shows that undergraduate student engagement changes from their first year to the fourth or fifth year in the university (Zhao et al., 2014). Especially for international students, although they seem to be engaged in academic courses, they are not as involved as domestic students in extracurricular activities. For this variable, there were four choices: 1st year, 2nd year, 3rd year, 4th year, 5+ years.
- *Major*: This independent variable is also very essential to the research design because the learning activities between undergraduates who are in STEM majors are different from students who are in liberal arts and fine arts majors. Current research has mainly investigated the effectiveness of gamified instructions for STEM students (Hamari et al., 2014; Kim, 2013; Xu, 2011). Although some researchers argue that educational gamification is more suitable for high-order thinking and problem-solving skills, it is still necessary to know about the needs of students who are from all majors for pleasurability in gamification. Only thinking from the designer's or the instructor's perspective might hinder research development. The choices of this variable were categorized by the list of school names at KU.

The third set of independent variables was students' technology/game backgrounds. This set was very important to this research because students' selections of technology-using frequency and their previous gamification experiences would influence the validity and the reliability of the research sample. Besides, previous research also shows the influence of users'

gaming habits on their gamification preferences (Landers & Callan, 2011). The criteria and choices for each variable in this set are listed as below:

- *Technology skill level*: Participants' technology skills would directly affect their user experience in educational gamification because they need to operate the activities on mobile devices or computers. Some research finds that nowadays undergraduates averagely spend 4.2 hours per week on computers for learning purposes (Krause, 2005). The mean score of this variable would decide whether the selected students in this research are special or as common as the sampling in other research. Students were given four choices: never, 1-3 hours per week, 4-6 hours per week, 7-9 hours per week. The definition of technology in this question was the electronic devices that have access to the Internet.
- *Gamification experiences*: By adding this variable, this research selected the sample of participants with gamification experiences to evaluate their pleasurability in learning environments. Participants chose from SNSs, massive multiplayer online games, online shopping, and massive online open courses (MOOCs) to indicate whether they have gamification experiences. For participants who chose "none of them" were excluded from the sample.
- *Gaming experience*: All participants were asked about their video/online game experiences in the survey. Some research found that users game experiences and their attitudes towards gamification were correlated given that the students who had more gaming experiences were more likely to have better gamification experiences (Landers & Callan, 2011). It is important for future designers and instructors to understand whether experienced gamers have as similar pleasurable experiences in

educational contexts as the students who did not play games. Four choices were provided to participants: never, 1-3 years, 4-6 years, 7+ years.

- *Gamer types*: Landers and Callan (2011) found that different gamer types might evaluate game factors differently. Some game players play for collecting experiences; some players play for accomplishments; and some players play for pure killing (story-driven). Participants were asked about their gaming purposes to find out what types of gamers they were. In the survey, the definition of each gamer type was provided to the participants in order to help participants understand the different categories.

Sampling

The participants in this research were the U.S. domestic and Chinese international undergraduate students at KU. 88 undergraduates participated in the pilot test; 279 undergraduates participated in the modified survey. This research recruited participants by distributing flyers and emailing acquaintances on campus. All participants in the pilot test and modified survey were full-time undergraduate students at KU.

KU has a large number of international students who come from more than 100 different countries around the world. Since 2010, the number of the international undergraduates has exceeded the graduate students at KU according to the statistics from its public website (https://iss.ku.edu/2010-2011). In 2016 spring semester, the enrollment of the international students composed 9.52% of the total student enrollment at KU. International undergraduate students shared 5.06% of the total student enrollment (https://iss.ku.edu/sites/iss.ku.edu/files/docs/Statistics/2016/S16%20-%20Total%20Enrollment.pdf). This trend of the international student enrollment at KU has provided a sufficient undergraduate student number base for conducting this research.

Participants provided their ratings for each pleasurable experience in educational contexts, as well as their demographic information, academic backgrounds, and gaming/technology backgrounds in the survey. The participants were further split into the group of U.S. domestic students and the group of Chinese international students for data comparisons.

Validity and Reliability

As this research was to design an instrument, validity and reliability were two very important indicators to measure the quality of this scale. Validity is used to measure whether the test truly measures what it claims. And reliability is used to measure whether the test can produce stable, consistent measurements.

Validity

In this research, validity referred to whether the Likert items included in the scale truly measure student senses of pleasurable learning experiences. External and internal validity should be investigated when testify the scale validity.

External Validity refers to the extent to which the results of a study are generalizable to other situations and to other people (Campbell & Stanley, 1966). The Likert items included in the surveys were selected from the shared pleasurable learning experiences from the target population. A pilot test and a modified survey were administrated to the KU undergraduates. Six items were excluded from the survey used in the pilot test and two new items were added in the modified survey. Based on the EFA results, the divisions of the sub-constructs under the surveys were similar with each other.

Internal Validity refers to a) the rigor with the study design, such as decisions concerning what was and was not measured, and b) the extent to which the research designer has taken into

account alternative explanations for any causal relationships they investigate (Huitt, 1998). Four types of internal validity should be considered to evaluate the instrument design:

Face Validity concerns about how a measure or procedure appears. In this research, the research procedure was to conduct qualitative inquiries to find out the shared pleasurable learning experiences; compare with PLEX categories; item review by an expert panel; distributed to target audience; run data analysis; name the subscales; map the subscales with the game factor based on the definitions of the each game factor and the items included in each subscale.

Criterion Related Validity refers to the accuracy of a measurement or procedure by comparing it with another measure or procedure that has been demonstrated to be valid. Two sub-categories of validity are under it: *Concurrent Validity* and *Predictive Validity*. *Concurrent Validity* means to have the criterion in the present; *Predictive Validity* means to have the criterion in the future.

Construct Validity refers to the theoretical supports for the test or procedure. *Convergent Validity* and *Discriminate Validity* are the two sub-categories under this validity. According to Carmines & Zeller (1991), three steps should be followed to understand the construct validity of research: "first, the theoretical relationships must be specified. Second, the empirical relationships between the measures of the concepts must be examined. And third, the empirical evidence must be interpreted in terms of how it clarifies the construct validity of the particular measure being tested" (Carmines & Zeller, 1991, p.23). PLEX was used to guide the design of Likert items in this research. This framework has been developed and experimented by Nokia research team and educators in gamification development (Arrasvuori et al., 2011; Kim, 2013).

Content Validity refers to the extent to which a measurement "reflects the specific intended domain of content" (Carmines & Zeller, 1991, p. 20). In this research, the Likert items

that were included in the survey could represent 13 PLEX categories and 2 non-PLEX categories. The exclusive PLEX categories were those feelings that were not suitable in the educational contexts.

Reliability

Reliability refers to the extent to which a measurement or procedure results in the same result on repeated trials. Reliability tests are important because the researchers would not be able to make claims about the generalizability of their research if they failed to prove the reliability of their measurements or procedures. There are four types of reliability:

Equivalency Reliability is the extent to which two items measure identical concepts at an identical level of difficulty. Correlation coefficients could be used to show the strength of the correlation between a dependent variable and one or more independent variables. The correlations between the 4 subscales extracted from EFA and the 9 predictors, e.g., age, gender, ethnicity, current studying year, major, technology skills, years of playing video/online/mobile games, frequencies of playing games per week, and gamer types, could be used to prove this type of reliability.

Stability Reliability is the consistency of measuring instruments over time. For this research, the subjects could be split into half and compare and correlate their ratings to prove this type of reliability.

Internal Consistency refers to the extent to which tests or procedures assess the same characteristic, skill or quality. This type of reliability could be analyzed by using Cronbach's alpha scale reliability tests. Based on the alpha scores, the researchers could understand the extent to which items on the scale focus on the same constructs.

Interater Reliability is the extent to which two or more raters agree. It addresses the consistency of the administration of a rating system. This type of reliability is not applicable for this research.

Statistical Analysis

A pilot test was administrated in order to testify the validity and reliability of this instrument. The Likert items used in the pilot test were based on the results in the interviews and reviewed by an expert panel for item clarifications. Undergraduate students who matched the criteria listed above were recruited to complete the online questionnaire. EFA was used in the pilot test, which indicated that four latent variables would be the best factor extraction for these Likert items. MANOVA was used in the pilot test to explore the explanation power of each independent variable on the dependent variables. One-way ANOVA was also conducted in order to compare the differences between the U.S. domestic undergraduates and Chinese international undergraduates on the four subscales.

Based on the results in the pilot test, six Likert items were excluded from the list of Likert scale. Focus groups were conducted and two new Likert items were added to the modified survey based on the EFA results in the pilot test. EFA was conducted in the modified survey in order to understand the underlying relationships among the new list of Likert items. The results in the statistical analysis of EFA were used to answer Research Question 1.

Then MANOVA was used to understand the main effect of each independent variable on the dependent variables. The main effect of each independent variable on the combination of all four factors that were extracted from EFA was analyzed. In order to further understand the unique contribution of each independent variable on each subscale, the results shown in the tests

for between-subject effects were also reported in the Results section. The design of measurement model for each factor is as below:

$$Factori \sim\sim Age + Gender + Ethnicity + CurrentStudyYear + Major + ComputerSkills + GamingYear + GamingFrequency + GamerType + e^{1}$$

The results from this statistical analysis were used to answer Research Question 2. As there was statistical significance in the result of MANOVA between *Ethnicity* and the subscales that were revealed from EFA, Bonferroni correction was used to adjust the level of significance. This adjustment was needed because of the scarce of research in studying cultural influences on gamification. Therefore, this research needed to be very conservative about the interpretation on the relationship between participants' *Ethnicity* and the *subscales*. The new level of significance was set as .006. But among the 9 categories under *Ethnicity*, only the population of Chinese and White participants was large. Therefore, a post-hoc test was conducted among *White*, *Chinese*, and *The Others* in order to balance the population. Bonferroni correction was also used to adjust the level of significance. The new level of significance for this post-hoc test was .017.

As the mean differences between Chinese and White participants were still statistically significant, one-way ANOVA was conducted to further compare the mean differences for undergraduate students' pleasurability in learning between Chinese international undergraduates and the U.S. domestic undergraduates. 119 U.S. domestic undergraduates were selected from the total participants to be compared with 81 Chinese international undergraduates. This statistical analysis was used to answer Research Question 2a.

[1] $Factor_i$ represents the mean score of each factor that was extracted from EFA for the modified survey. e represents the errors in the measurement.

Results

This chapter provides the research results, which includes the descriptive statistics about the subjects that participated in the development of the instrument and statistical analysis of the collected data. The instrument developed in this research was a 4-point Likert agree/disagree scale (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree) with the purpose of understanding undergraduate student preferences for the pleasurable learning experiences. 22 common pleasurable learning experiences were analyzed in the finalized instrument (See Appendix K). This Likert scale was distributed to 279 undergraduates recruited by email and social media tools at the University of Kansas (KU) via Qualtrics. 12 demographic questions were also included in the survey in order to further analyze participants' attitudes towards the pleasurable learning experiences (See Appendix L).

Findings of the qualitative research activities and a pilot test are also reported in this chapter. Interviews and focus groups were conducted in order to generate genuine descriptions of pleasurable learning experiences among the undergraduate students. One pilot test with 88 participants was conducted in order to develop a more reliable instrument.

Findings of the Qualitative Research Methods

Interviews

7 interviews were conducted at KU, which included 4 undergraduate students and 3 GTAs. These interviewees came from different majors, such as education, engineering, humanities & international studies, and social & behavioral sciences. 3 interviewees were female and 4 interviewees were male. The recordings of the interviews were transcribed and saved in a password-secured personal computer. The transcripts have been sent back to the interviewees for member checks before using them for generating the survey items.

After reading the transcripts of the interviews and comparing with PLEX categories, 34 statements of pleasurable learning experiences were generated and reviewed by an expert panel. These statements were chosen from the interview conversations because they were mentioned by at least two interviewees and were relevant to the PLEX categories. Table 3 is the list of the survey items generated based on the interviews. This list of survey items was used in the pilot test.

Table 3

Matrix of 34 Survey Items and Pleasurable Learning Experiences (PLLEX) Categories

Survey Item	PLLEX Categories
1. I like taking notes during instructions.	Expression
2. I like instructors to use PowerPoint slides.	Visualization*
3. I like instructors to write the contents on the white/black board.	Visualization*
4. I like learning new contents outside of the class while doing activities in class.	Interaction*
5. I like instructors to use interactive technology (e.g., i-clickers and social media tools).	Interaction*
6. I like instructors to share class materials with us.	Control
7. I like previewing the new contents.	Exploration
8. I like trying new things.	Discovery
9. I like challenges.	Challenge
10. I like instructors to give us guidance/directions when we are doing in-class activities.	Submission
11. I like receiving instructors' comments on my assignments.	Fellowship
12. I like participating in virtual discussion groups, e.g., discussion board on Blackboard application.	Expression
13. I like being able to check my progress.	Control
14. I prefer subjective assessments, such as essays and short answers.	Expression
15. I prefer objective assessments, such as multiple-choice questions and tests.	Expression
16. I like getting quick feedback on my performances from the instructors.	Fellowship
17. I like reviewing contents.	Control
18. I like collaborating with my classmates.	Fellowship/ Interaction*

19. I like learning new materials that are related to my prior knowledge.	Control
20. I like self-paced learning experiences.	Control/Relaxation
21. I am self-motivated.	Control
22. I like interactive in-class activities.	Interaction*
23. I like learning knowledge and skills that can apply directly to the real world problems.	Exploration
24. I like instructors to provide clear rubrics to us.	Control
25. I like instructors to provide clear syllabus to us.	Control
26. I like learning knowledge and skills that are related to the final exam of the course.	Control
27. I like instructors' attentions.	Fellowship
28. I like building personal connections with instructors.	Fellowship
29. I like instructors being accessible.	Control/Fellowship
30. I like knowing other students' grades on the assignments.	Competition
31. I like the feeling of completing a major task.	Completion
32. I like instructors to use scenarios during instructions.	Fantasy/ Sensation
33. I like instructors to use funny pictures and relevant quotations during instructions.	Humor
34. I like instructors to make jokes and tell stories during instruction.	Humor

Note. Categories with "*" are non-PLEX categories, which are pleasurable experiences particularly in learning environments.

Several original PLEX categories have been deleted because they were not suitable in academic learning environments and were not mentioned by the interviewees: Captivation, Cruelty, Eroticism, Nurture, Simulation, Subversion, Suffering, Sympathy, and Thrill. Two new categories of pleasurable learning experiences were found based on the interviews: Visualization and Interaction. Visualization was defined as recalling or forming mental images. Interaction was defined as a reciprocal action, effect, or influence. The new set of PLEX categories and non-PLEX categories was named as pleasurable learning experiences (PLLEXs).

Focus Groups

The purposes of conducting focus groups were a) to explore if there were more pleasurable learning experiences as participants might find more pleasurable learning experiences from this brainstorming activity; b) to compare if there were any difference between Chinese international students and U.S. domestic students in their preferences for pleasurability in learning. Two focus groups were conducted with the students at KU. One group consisted of 4 Chinese international students; 2 of them were undergraduates and 2 of them were GTAs. The other focus group consisted of 3 U.S. domestic students; 2 were GTAs and 1 was an undergraduate student.

The participants in the focus groups wrote down the pleasurable learning experiences by themselves at the beginning and then shared their experiences with the other participants in the groups. From the discussions, participants frequently agreed with others' pleasurable learning experiences and recalled the pleasurable learning experiences that they observed in classes. At the end of the focus groups, participants discussed about the 34 found pleasurable learning experiences in the interviews and added two new experiences listed as below to the list. No significant difference was found when these two groups of students described their pleasurable learning experiences:

1. I like instructors to use videos during instructions.
2. I like building personal connections with my classmates.

Findings of the Pilot Test

A pilot test was conducted in 2015 summer. The survey items included in this pilot test were the 34 pleasurable learning experiences generated based on the interviews only due to the time limitation. Focus groups were not conducted until 2015 winter. 88 KU undergraduate students participated in this pilot test, which were recruited by on-campus flyers and social

media tools. Table 4 – 6 reveal the descriptive statistics of the subjects that participated in the pilot test:

Table 4

Demographic Information of the Subjects in the Pilot Test

	Age (N = 88)					Gender (N = 88)	
	18-19	20-21	22-23	24-25	26+	Male	Female
Frequency	7	44	26	8	3	42	46
%	8	50	29.5	9.1	3.4	47.7	52.3
	Ethnicity (N = 88)						
	White	Hispanic	Black	Chinese	Other*		
Frequency	52	6	3	23	4		
%	59.1	6.8	3.4	26.1	4.5		

Note. Subjects who chose "Other" in *Ethnicity* indicated themselves as multi-racial.

Table 5

Academic Background of the Subjects in the Pilot Test

	Academic Year (N = 88)				
	1st	2nd	3rd	4th	5th +
Frequency	4	9	28	47	15
%	4.5	10.2	31.8	36.4	17

Major	Frequency	%
Architecture	2	2.3
Arts	3	3.4
Business	12	13.6
Education	6	6.8
Engineering	16	18.2
Health Professions	6	6.8
Health, Sport, and Exercise Sciences	7	8.0
Humanities & International Studies	1	1.1
Journalism & Mass Communications	2	2.3
Law (Pre-Law)	2	2.3
Medicine (Pre-Med)	3	3.4
Natural Science & Math	4	4.5
Nursing	3	3.4
Pharmacy	1	1.1

Public Affairs & Administration	2	2.3
Social & Behavioral Sciences	16	18.2
Social Welfare	1	1.1
Undecided	1	1.1
Total	88	100.0

Table 6

Gaming Background of the Subjects in the Pilot Test

	Years of Playing Video/Online/Mobile Games			
	1-3 years	4-6 years	7+ years	Never
Frequency	19	14	44	11
%	21.6	15.9	50	12.5
	Frequency of playing video/online/mobile games per week			
	1-5 hours	6-10 hours	11-15 hours	16+ hours
Frequency	48	15	5	3
%	54.5	17	5.7	3.4
	Gamer Type			
	Story-driven Solo Gamer	Social Gamer	Solo-limited Gamer	Hardcore Online Gamer
Frequency	11	25	6	5
%	12.5	28.4	6.8	5.7

Skewness and Kurtosis of the 34 survey items were within the range of ±1.96, which indicated that these survey items were normally distributed (George & Mallery, 2010; See Appendix C). It is important to check the normality tests of the survey items because one of the prerequisites of conducting factor analysis is that the items should be normally distributed. Then exploratory factor analysis (EFA) was conducted to further study the underlying relationships among these 34 survey items. 6 survey items were excluded from the scale along the factor extraction as they had respectively low communalities. Communalities indicate the amount of variance in each variable that is accounted for by the common factors. If the communality were very low, it means that the variable does not fit the construct well. 0.30 was used as the cut-off

score to suppress the survey items that held low communalities in the pilot test. The eliminated survey items were:

3. I like instructors to write the contents on the white/black board.

15. I like assessments as essays and short answers.

16. I prefer objective assessments, such as multiple-choice questions and tests.

20. I like self-paced learning experiences.

23. I like learning knowledge and skills that can apply directly to the real world problems.

30. I like knowing other students' grades on the assignments.

4 factors were revealed by using principal component analysis (PCA) as the extraction method, which is a widely used method for factor extraction in EFA (Polit, 2012; See Table 7). The cut-off factor loading score was .40, which is a common cut-off score in EFA (Stevens, 1992). Two items, e.g., item 4 and item 13, had cross-loadings but did not affect the internal consistencies of the factors. Then the factor model was rotated by an oblique rotation method Promax for analysis because the component correlations were higher than .32 (Brown, 2009; See Appendix D). The rotations are necessary in factor analysis because they help simplify the structure and produces more interpretable factors. Two types of rotation methods can be used in factor analysis. One type is called oblique rotation and it can be used with correlated factors; the other type is called orthogonal rotation and it can be used with uncorrelated factors. Most research believes that factors are correlated when their correlation scores are higher than .32. Promax is one of the most popular oblique rotation methods.

Table 7

Survey Items under Each Factor in the Pilot Test

	Factor Loadings			
Survey Items	Factor 1	Factor 2	Factor 3	Factor 4

2. I like instructors to use PowerPoint slides.	0.81		
1. I like taking notes during instructions.	0.73		
10. I like instructors to give us guidance/directions when we are doing in-class activities.	0.64		
17. I like reviewing contents.	0.62		
26. I like learning knowledge and skills that are related to the final exam of the course.	0.56		
25. I like instructors to provide clear syllabus to us.	0.55		
6. I like instructors to share class materials.	0.51		
13. I like being able to check my progress.	0.50	0.47	
7. I like previewing new content.	0.49		
34. I like instructors to make jokes and tell stories during instruction.		0.91	
33. I like instructors to use funny pictures and relevant quotations during instructions.		0.83	
16. I like getting quick feedback on my performances from the instructors.		0.66	
31. I like the feeling of completing a major task.		0.62	
32. I like instructors to use scenarios during instructions.		0.55	
24. I like instructors to provide clear rubrics to us.		0.49	
11. I like receiving instructors' comments on my assignments.		0.42	
9. I like challenges.			0.79
28. I like building personal connections with instructors.			0.78
8. I like trying new things.			0.62
21. I am self-motivated.			0.59
29. I like instructors being accessible.			0.53
27. I like instructors' attentions.			0.53
19. I like learning new materials that are related to my prior knowledge.			0.44

12. I like participating in virtual discussion groups, e.g., discussion board on Blackboard application.		0.71
5. I like instructors to use interactive technology (e.g., i-clickers and social media tools).		0.68
22. I like interactive in-class activities.		0.63
18. I like collaborating with my classmates.		0.63
4. I like learning new contents outside of the class while doing activities in class.	0.52	0.53

Note. Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalization.

Researchers often use eigenvalues to decide the number of factors as the eigenvalues indicate the amount of variance in the overall scale that can be accounted by each factor. Although 8 factors had eigenvalues above 1 (See Figure 5), some components consisted of fewer than 4 survey items. Since too few survey items within a substantial scale would affect the internal consistency of the respective scale, this pilot test opted to include the first 4 factors that explained 50.53% of the total variance (See Table 8). Beaver et al. (2013) indicated that the acceptable amount of explained variance in EFA could be as low as 50%.

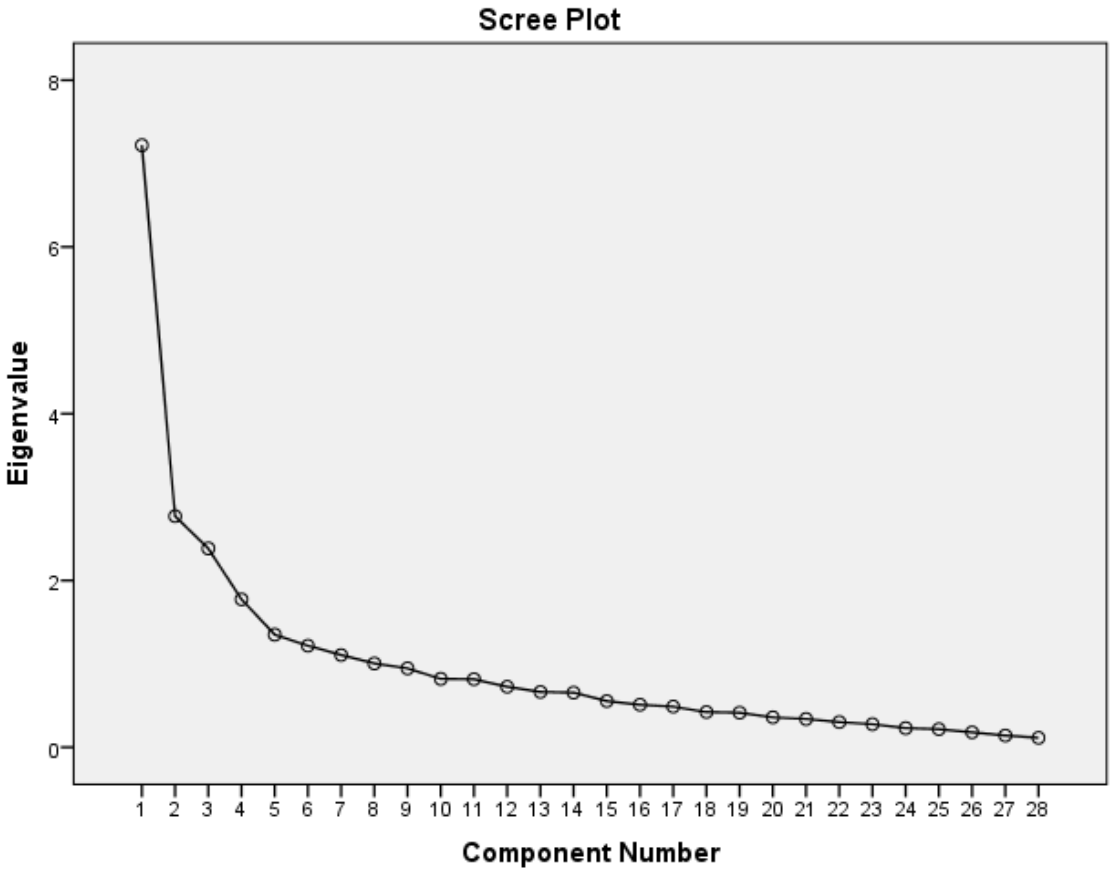

Figure 5. Scree Plot of Factors in EFA from the Pilot Test

Table 8

Total Variance Explained by the Extracted Factors in the Pilot Test

	Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %
1	7.22	25.79	25.79
2	2.77	9.90	35.68
3	2.38	8.52	44.20
4	1.77	6.33	50.53

Cronbach's alpha scale reliability tests were conducted to analyze the internal consistencies of the remaining survey items in each substantial scale (See Appendix E).

Cronbach's alpha is a measure of internal consistency, which is used to inspect how closely correlated a set of items are as a group. It is considered to be a measure of scale reliability, particularly in the measure of internal consistencies in substantial constructs. The acceptable Cronbach's alpha range is from .70 to .90 (Tavakol & Dennick, 2011). A low value of Cronbach's alpha would indicate low interrelatedness between items, and a high value of Cronbach's alpha would suggest the redundancies of items. The reliability test scores for the factors were .84, .82, .76, and .69 as shown in Appendix E. Given the reliability test score of factor 4 was below .70, more survey items should be generated to increase its internal consistency (George & Mallery, 2003).

Findings of MANOVA in the pilot test

MANOVA was conducted to explore the influential independent variables (IVs) that contributed to the variance in undergraduate student preferences for pleasurable learning experiences. Table 9 shows that 9 IVs were included in the overall model. The 4 dependent variables (DVs) were the mean score of each factor that was extracted from EFA. *Ethnicity* was the only IV that yielded statistically significant differences among participants' preferences with $\Lambda = .42$, $F(16, 168) = 2.32$, $p < .01$, $R^2 = .19$.

Table 9

Tests of IVs that Cause Statistically Significant Differences among the DVs in the Pilot Test

IV	Wilks' λ	F	df1	df2	*p*	η^2
Age	0.64	1.17	16	120	.299	.105
Gender	0.88	1.281[b]	4	39	.294	.116
Ethnicity	0.42	2.42	16	120	.003**	.193
Current Study Year	0.71	0.88	16	120	.599	.081
Major	0.27	0.91	68	155	.661	.281
Frequency of Using Computers for Learning per Week	0.68	1.36	12	103	.200	.121
Years of Playing Online/Video/Mobile Games	0.68	1.38	12	103	.188	.123

Frequency of Playing Online/Video/Mobile Games per Week	0.70	0.92	16	120	.545	.085
Gamer Type	0.63	0.97	20	130	.500	.109

Note. IV = independent variable. Mean difference is considered to be statistically significant at the significance level of .05. *$p < .05$, ** $p < .01$, *** $p < .001$.

As *Ethnicity* yielded statistically significant differences in the overall scale, the results shown in the tests of between-subjects effects (See Appendix F) were analyzed to further understand the main effect of *Ethnicity* on each Factor when the other IVs were present in the measurement model. Table 10 shows that *Ethnicity* resulted in statistically significant differences in factor 1 with $R^2 = .37$, $F(4, 86) = 5.94$, $p < .01$ and factor 4 with $R^2 = .23$, $F(4, 86) = 3.03$, $p < .05$. However, with the presences of all other IVs in the model, no factor had statistically significant between-subjects differences at the significance level of .05 (See Appendix F). This result indicates that the main effect of *Ethnicity* on each factor has been suppressed by the combinations of other IVs.

Table 10

Effect of Ethnicity on Each Factor in the Pilot Test

IV	DV	df	F	*p*.	η^2
Ethnicity	Factor 1	4	5.94	.001**	.367
	Factor 2	4	1.16	.344	.101
	Factor 3	4	1.36	.264	.117
	Factor 4	4	3.03	.028*	.228

Note. IV = independent variable. DV = dependent variable. Using level of significance is $p < .05$. * $p < .05$, ** $p < .01$, *** $p < .001$.

Findings of the Differences between Chinese and U.S. Undergraduates in the Pilot Test

As *Ethnicity* yielded a statistical significance among different groups of students, further comparisons have been conducted between Chinese international undergraduates and the U.S. domestic undergraduates because the Asian group in this pilot test was Chinese international

students and the majority of the white participants were the U.S. domestic undergraduates. The descriptive statistics of these two groups of participants are shown in Table 11 – 13.

Table 11

Demographic Information of the Chinese International Undergraduates and U.S. Domestic Undergraduates in the Pilot Test

	Age					Gender	
	18-19	20-21	22-23	24-25	26+	Male	Female
Chinese (N = 23)	2	9	5	5	2	13	10
%	8.7	39.1	21.7	21.7	8.7	56.5	43.5
US (N = 62)	5	33	20	3	1	10	36
%	8.1	53.2	32.3	4.8	1.6	41.9	58.1

Table 12

Academic Information of the Chinese International Undergraduates and U.S. Domestic Undergraduates in the Pilot Test

	Academic Year				
	1st	2nd	3rd	4th	5th +
Chinese (N = 23)	1	5	6	7	4
%	4.3	21.7	26.1	30.4	17.4
US (N = 62)	3	4	20	24	11
%	4.8	6.5	32.3	38.7	17.7

Major	Chinese	%	United States	%
Architecture	2	8.7	3	4.8
Arts				
Business	5	21.7	6	9.7
Education	3	13.0	3	4.8
Engineering	9	39.1	7	11.3
Health Professions			6	9.7
Health, Sport, and Exercise Sciences			7	11.3
Humanities & International Studies			1	1.6
Journalism & Mass Communications			2	3.2
Law (Pre-Law)			2	3.2
Medicine (Pre-Med)	1	4.3	2	3.2
Natural Science & Math			3	4.8
Nursing			3	4.8
Pharmacy	1	4.3		
Public Affairs & Administration			2	3.2
Social & Behavioral Sciences	1	4.3	14	22.6
Social Welfare			1	1.6

Undecided	1	4.3		
Total	23	100	62	100

Table 13

Gamification Experiences of the *Chinese International Undergraduates and U.S. Domestic Undergraduates in the Pilot Test*

	Frequency of Using Computer for Learning per Week			
	1-3 hours/week	4-6 hours/week	7-9 hours/week	10+ hours/week
Chinese (N = 23)	0	4	4	15
%	0	17.4	17.4	65.2
US (N = 62)	6	26	14	16
%	9.7	41.9	22.6	25.8

	Gamification Experiences			
	Massive Multiplayer Online Games	Massive Open Online Courses	Online Shopping	Social Network Sites
Chinese (N = 23)	10	13	20	20
%	43.5	56.5	87.0	87.0
US (N = 62)	28	23	61	61
%	45.2	37.1	98.4	98.4

	Years of Playing Video/Online/Mobile Games			
	1-3 years	4-6 years	7+ years	Never
Chinese (N = 23)	10	7	5	1
%	43.5	30.4	21.7	4.3
US (N = 62)	8	7	37	10
%	12.9	11.3	59.7	16.1

	Frequency of Playing Video/Online/Mobile Games per Week				
	1-5 hours	6-10 hours	11-15 hours	16+ hours	Never
Chinese (N = 23)	11	6	3	1	2
%	47.8	26.1	13.0	4.3	8.7
US (N = 62)	35	8	2	2	15
%	56.5	12.9	3.2	3.2	24.2

	Gamer Type					
	Story-driven Solo Gamer	Social Gamer	Solo-limited Gamer	Hardcore Online Gamer	Control/identity Solo Gamer	Casual Gamer
Chinese (N = 23)	5	5	0	2	4	7
%	21.7	21.7	0	8.7	17.4	30.4
US (N = 62)	5	18	6	3	5	24
%	8.2	29.5	9.8	4.9	8.2	39.3

One-way ANOVA was conducted to compare the group differences between Chinese international undergraduates' and U.S. domestic undergraduate students' attitudes towards pleasurable learning experiences. The IV was the group membership, Chinese vs. U.S.; and the DVs were the group means of those 4 factors extracted from EFA. Table 14 shows that these two groups of undergraduates differ statistical significantly in their preferences for factor 1 with $F(1, 86) = 6.55$, $p < .05$; factor 2 with $F(1, 86) = 9.47$, $p < .01$; and factor 3 with $F(1, 86) = 10.77$, $p < .01$. The mean scores of Chinese international undergraduates for these 3 factors were consistently lower than the U.S. domestic undergraduates.

Table 14

Comparisons between Chinese International Undergraduates and U.S. Domestic Undergraduates

Source	Group	n	M	SD	F(1, 86)	*p*
Factor1	US	65	3.56	0.44	6.55	.012*
	CHINA	23	3.29	0.41		
	Total	88	3.49	0.45		
Factor2	US	65	3.45	0.37	9.47	.003**
	CHINA	23	3.16	0.44		
	Total	88	3.37	0.41		
Factor3	US	65	3.44	0.36	10.77	.001***
	CHINA	23	3.15	0.36		
	Total	88	3.36	0.38		
Factor4	US	65	2.80	0.54	1.96	.165
	CHINA	23	2.97	0.42		
	Total	88	2.85	0.52		

Note. n = number of participants. M = means. SD = standard deviation. Significant at the .05 level. *$p < .05$, **$p < .01$, ***$p < .001$.

Findings of the Modified Survey

The population of the pilot test was too small, and the survey items were purely based on the interviews. From the results of the pilot test, the statistics suggested that more items should

be added to factor 4 in order to improve its internal reliability. A new survey was distributed to the undergraduate students at KU with the purpose of further testing the validity and reliability of the modified scale.

The survey items in this modified survey consisted of the 28 remaining survey items from the pilot test and 2 new items from the focus groups. The 2 new items were added into the scale as a result of focus groups and the results of the pilot test. In particular, survey item 30 was used to increase the internal consistency of factor 4 from the pilot test results because this survey item was a PLLEX that was related to student attitudes towards interactions and collaborations as the rest items in factor 4. Table 15 shows the matrix of the new survey items and the PLLEX categories. Based on the EFA results from the pilot test, more PLEX categories were excluded from this new survey: Competition, Exploration, and Relaxation.

Table 15

Matrix of the Survey Items and PLLEX Categories in the Modified Survey

Survey Item	PLLEX Categories
1. I like taking notes during instructions.	Expression
2. I like instructors to use PowerPoint slides.	Visualization*
3. I like learning new contents outside of the class while doing activities in class.	Interaction*
4. I like instructors to use interactive technology (e.g., i-clickers and social media tools).	Interaction*
5. I like instructors to share class materials with us.	Control
6. I like previewing the new contents.	Exploration
7. I like trying new things.	Discovery
8. I like challenges.	Challenge
9. I like instructors to give us guidance/directions when we are doing in-class activities.	Submission
10. I like receiving instructors' comments on my assignments.	Control/Submission
11. I like participating in virtual discussion groups, e.g., discussion board on Blackboard application.	Expression
12. I like being able to check my progress.	Control

13. I like getting quick feedback on my performances from the instructors.	Control/Submission
14. I like reviewing contents.	Control
15. I like collaborating with my classmates.	Fellowship/Interaction*
16. I like learning new materials that are related to my prior knowledge.	Control
17. I am self-motivated.	Control
18. I like interactive in-class activities.	Interaction*
19. I like instructors to provide clear rubrics to us.	Control
20. I like instructors to provide clear syllabus to us.	Control
21. I like learning knowledge and skills that are related to the final exam of the course.	Control
22. I like instructors' attentions.	Fellowship
23. I like building personal connections with instructors.	Fellowship
24. I like instructors being accessible.	Control/Fellowship
25. I like the feeling of completing a major task.	Completion
26. I like instructors to use scenarios during instructions.	Fantasy/Sensation
27. I like instructors to use funny pictures and relevant quotations during instructions.	Humor
28. I like instructors to make jokes and tell stories during instruction.	Humor
29. I like instructors to use videos during instructions.	Sensation/Visualization*
30. I like building personal connections with my classmates.	Fellowship/Interaction*

Note. Categories with "*" are non-PLEX categories, which are pleasurable experiences particularly in learning environments.

283 undergraduates completed this modified survey, but 4 respondents were excluded from the data analysis as their responses for survey item 26 resulted in a positive skewed distribution. After taking out the outliers, the valid sample size used in this research was 279. Table 16 – 18 show the descriptive statistics of the subjects that participated in this modified survey.

Table 16

Demographic Information of the Subjects in the Modified Survey

	Age				Gender	
	18-19	20-21	22-23	24+	Male	Female

Frequency	74	109	61	35	122	157
%	26.5	39.1	21.9	12.5	43.7	56.3

	Hispanic, Latino, Latino Origin			Are You an International Student in the U.S.	
	No	Mexcian, Mexican Am., Spanish Origin	Another Hispanic, Latino, or Spanish Origin	Yes	No
Frequency	267	10	2	84	195
%	95.7	3.6	0.8	30.1	69.9

	Ethnicity			
	White	Black	American Indian or Alaska Native	Chinese
Frequency	171	9	7	82
%	61.3	3.2	2.5	29.4

Note. The racial categories in Ethnicity were adapted from US Census 2010. Some racial groups with very few subjects are not included in the table. 1 subject was Asian Indian; 1 subject was Filipino; 1 subject was Japanese; 1 subject was Korean; 1 subject was Vietnamese; 1 subject indicated that himself was multi-racial; and 4 missing data. These 4 missing data were those who selected themselves with Mexican, Mexican Am., Spanish Origins.

Table 17

Academic Background of the Subjects in the Modified Survey

	Academic Year				
	1st	2nd	3rd	4th	5th +
Frequency	55	58	69	66	31
%	19.7	20.8	24.7	23.7	11.1

Major	Frequency	%
Architecture	13	4.7
Arts	23	8.2
Business	51	18.3
Education	24	8.6
Engineering	34	12.2
Health Professions	16	5.7
Health, Sport, and Exercise Sciences	14	5.0
Humanities & International Studies	12	4.3
Journalism & Mass Communications	9	3.2
Law (Pre-Law)	5	1.8
Medicine (Pre-Med)	10	3.6
Music	3	1.1
Natural Science & Math	10	3.6
Nursing	4	1.4

Pharmacy	1	.4
Public Affairs & Administration	2	.7
Social & Behavioral Sciences	38	13.6
Social Welfare	2	.7
Undecided	8	2.9
Total	279	100.0

Table 18

Gamification Background of the Subjects in the Modified Survey

	Frequency of Using Computer for Learning per Week				
	Never	1-3 Hours/Week	4-6 Hours/Week	7-9 Hours/Week	10+ Hours/Week
Frequency	2	22	77	68	110
%	0.7	7.9	27.6	24.4	39.4
	Gamification Experiences				
	Massive Multiplayer Online Games		Massive Open Online Courses	Online Shopping	Social Network Sites
Frequency	120		96	255	249
%	43		34.4	91.4	89.2
	Years of Playing Video/Online/Mobile Games				
	1-3 years		4-6 years	7+ years	Never
Frequency	55		39	131	54
%	19.7		14	47	19.4
	Frequency of Playing Video/Online/Mobile Games per Week				
	1-5 hours	6-10 hours	11-15 hours	16+ hours	Never
Frequency	137	39	8	14	81
%	49.1	14	2.9	5	29
	Gamer Type*				
	Story-driven Solo Gamer	Social Gamer	Solo-limited Gamer	Hardcore Online Gamer	Control/identity Solo Gamer	Casual Gamer
Frequency	32	52	26	10	30	72
%	11.5	18.6	9.3	3.6	10.8	25.8

Note. *56 subjects indicated that they did not belong to any gamer type listed above, which were 20.10% among the total population.

29 among the 30 survey items were normally distributed by checking their Skewness and Kurtosis statistics. Item 26 was not normally distributed as its Kurtosis statistic exceeded 1.96. By checking the standardized values of survey item 26, 4 respondents were excluded from the sample. Then all of the survey items were normally distributed.

Then a Cronbach's alpha scale reliability test was conducted to analyze the correlations between each survey item and the overall scale. 2 items, e.g., item 1 and item 11, were excluded from the scale because their low correlations with the rest of the items. The results from Cronbach's alpha scale reliability test indicated that when these two items were deleted, the internal consistency of the overall scale would increase (See Appendix G).

Findings of EFA in the Modified Survey

After the item analysis, EFA was used to investigate the underlying relationships among the remaining 28 survey items. PCA was used to extract the factors and then rotated by the oblique rotation method Promax because the factor correlations were above .32 (See Table 19).

Table 19

Factor Correlation Matrix of the Modified Survey

Factor	1	2	3	4
1	1.00	0.35	0.18	0.54
2	0.35	1.00	0.28	0.31
3	0.18	0.28	1.00	0.36
4	0.54	0.31	0.36	1.00

Note. Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalization.

In the process of factor extractions, 5 survey items were eliminated from the scale because of their lower communalities:

2. I like it when instructors make their PowerPoint slides available to me.
3. I like learning new content outside of the class while doing activities in class
6. I like previewing new content.
16. I like learning new materials that is related to my prior knowledge.
29. I like the feeling of completing a major task.

Although 5 factors had eigenvalues above 1 (See Figure 6), one of the factors had fewer than 4 survey items because one of the survey items had factor loading lower than .40. Thus, for the parsimony's sake, 4 factors was the best number of factor extraction. 54.35% of variance in the overall scale could be explained by these 4 factors as shown in Table 20.

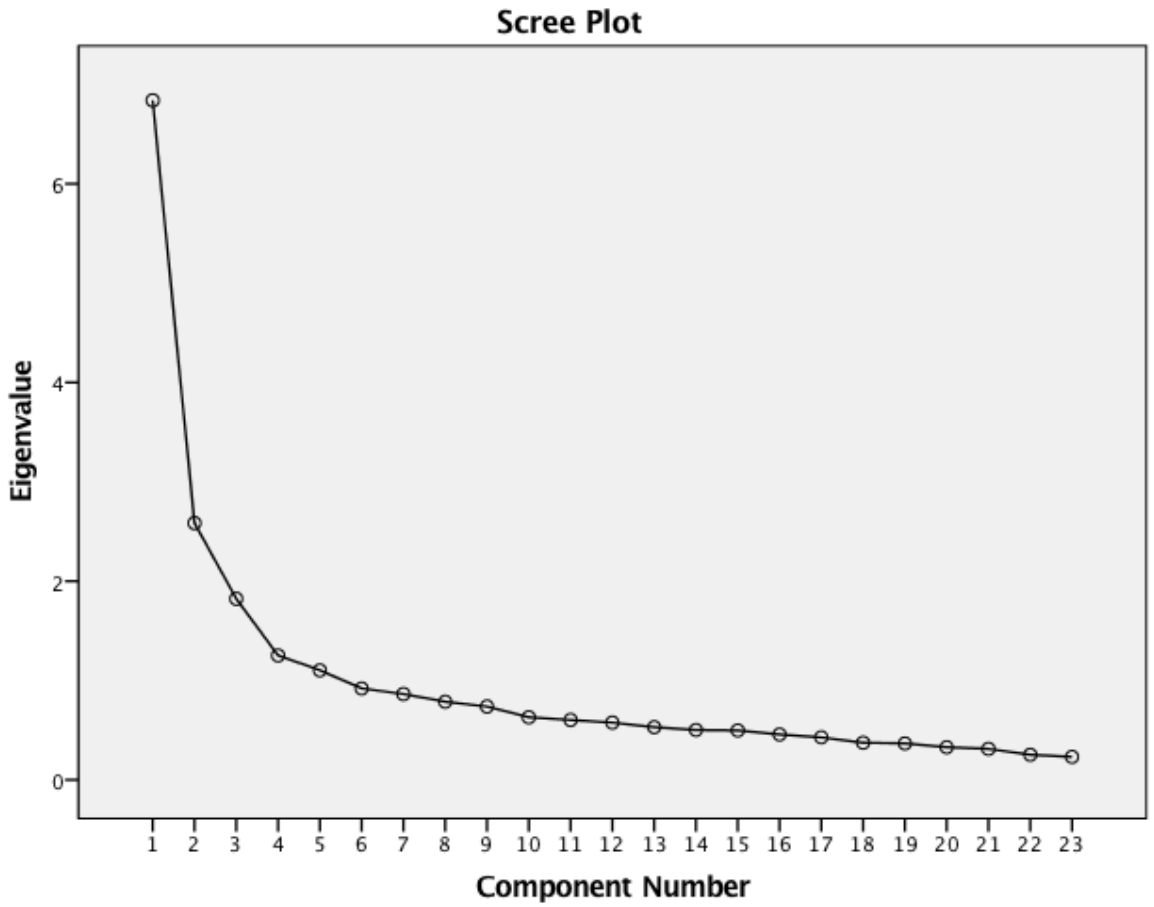

Figure 6. Scree Plot of the Factors in the Modified Survey

Table 20

Total Variance Explained by the 4 Factors in the Modified Survey

	Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %
1	6.84	29.74	29.74
2	2.59	11.24	40.98
3	1.82	7.93	48.90
4	1.25	5.44	54.35

Based on the description of each survey item, the 4 factors can be categorized as below (See Appendix H):

- Factor 1 consisted of 8 items, which were related to undergraduates' *Preference for Instruction*;
- Factor 2 consisted of 4 items, which were related to undergraduates' *Preference for Teaching Style*;
- Factor 3 consisted of 6 items, which were related to undergraduates' *Preference for Activities*;
- Factor 4 consisted of 5 items, which were related to undergraduates' *Preferences for Learning Effectiveness*.

The cut-off factor loading score was .40 (Stevens, 1992). One cross-loading was found with survey item 8, but it did not affect the internal consistency of factor 3 by checking the interrelation of this item with the rest items in factor 3 (See Table 21).

Table 21

Cronbach's Alpha Scale Reliability Tests of the Substantial Scales in the Modified Survey

Substantial Scale	Cronbach's Alpha	N of Items
Factor 1	.86	8
Factor 2	.84	4
Factor 3	.75	6
Factor 4	.74	5

Each factor obtained a Cronbach's alpha statistic higher than .70, which suggested that these 4 subscales were quite reliable. However, survey item 26 in factor 2 had a respectively lower correlation with the rest items in factor 2. Its Cronbach's alpha scale reliability test indicated that if this item were deleted, the scale reliability statistic would increase from .84

to .86. Therefore, item 26 was excluded from factor 2. The other survey items all fitted well within each factor.

Findings of MANOVA in the Modified Survey

The mean scores of the 4 factors that were revealed from EFA have been computed as the new DVs for each respondent in order to compare if there was any statistically significant difference caused by the IVs. The equation of each DV is as below:

- Factor 1 = (Item 5 + Item 9 + Item 10 + Item 12 + Item 13 + Item 19 + Item 20 + Item 21) / 8
- Factor 2 = (Item 27 + Item 28 + Item 29) / 3
- Factor 3 = (Item 4 + Item 7 + Item 8 + Item 15 + Item 18 + Item 30) / 6
- Factor 4 = (Item 14 + Item 17 + Item 22 + Item 23 + Item 24) / 5

Overall model

MANOVA was used to analyze the main effects of the 9 IVs on the 4 DVs. *Age*, *Gender*, *Ethnicity*, *Current Study Year*, *Major*, *Frequencies of Using Computers for Learning per Week*, *Years of Playing Online/Video/Mobile Games*, *Frequencies of Playing Online/Video/Mobile Games per Week*, and *Gamer Types* were the 9 IVs used in the measurement model to investigate their main effects on the 4 DVs.

Table 22 indicates that *Gender* ($\eta^2 = .11$, $\Lambda = .89$, $F(4, 217) = 6.41$, $p < .001$), *Ethnicity* ($\eta^2 = .07$, $\Lambda = .76$, $F(36, 814) = 1.74$, $p < .01$), and *Frequencies of Using Computers for Learning per Week* ($\eta^2 = .04$, $\Lambda = .83$, $F(16, 664) = 2.53$, $p < .01$) were the IVs that resulted in statistically significant differences in the participants' agreements with the 4 factors.

Table 22

Tests of IVs that Cause Statistically Significant Differences among the DVs in the Modified Survey

IV	Wilks' Λ	F	df1	df2	*p*.	η^2
Age	0.95	1.02	12	574	.430	0.02
Gender	0.89	6.41	4	217	.000***	0.11
Ethnicity	0.76	1.74	36	815	.005**	0.07
Current Study Year	0.96	0.60	16	664	.887	0.01
Major	0.68	1.21	72	856	.124	0.09
Frequency of Using Computers for Learning per Week	0.83	2.54	16	664	.001**	0.04
Years of Playing Online/Video/Mobile Games	0.95	0.92	12	574	.529	0.02
Frequency of Playing Online/Video/Mobile Games per Week	0.92	1.10	16	664	.348	0.02
Gamer Type	0.88	1.22	24	758	.213	0.03

Note. IV = independent variable. Mean difference is considered to be statistically significant at the significance level of .05. *$p < .05$, ** $p < .01$, *** $p < .001$.

Table 23 shows the MANOVA results of IV effects of between-subjects on each factor extracted from EFA in the modified survey (See Appendix I). The results would tell whether there were group differences in each factor caused by these 9 IVs. The statistics of this table show that these 279 undergraduates differed statistical significantly in their agreements with factor 1 ($R^2 = .32$, $F(52, 220) = 1.95$, $p < .001$), factor 3 ($R^2 = .31$, $F(52, 220) = 1.87$, $p < .01$), and factor 4 ($R^2 = .32$, $F(52, 220) = 1.98$, $p < .001$).

Table 23

Analysis of Variance in Each Factor from EFA in the Modified Survey

Source	DV	df1	df2	F	*p*	η^2
Corrected	Factor 1	52	220	1.95	.000*	.32
Model	Factor 2	52	220	1.37	.064	.24
	Factor 3	52	220	1.87	.001*	.31
	Factor 4	52	220	1.98	.000*	.32

Note. DV = dependent variable. The significance level is adjusted by using Bonferroni correction at .001. *$p < .001$.

As there were very few participants in the ethnical groups other than the Chinese and White participants, post-hoc tests among White, Chinese, and the other ethnical groups with each

factor were conducted in order to understand if there were any group difference and to control Type I error (See Appendix J). Bonferroni correction was used to compare the group differences. The new significance level was .017. Statistically significant group differences were found between White and Chinese participants in factor 1 and factor 4 with $p < .001$. Chinese and the other ethnical groups also differed statistical significantly in their agreements with factor 4 with $p < .01$.

Each subscale

As there were between-subjects differences in factor 1, factor 3, and factor 4, more statistics were read to find out which IVs resulted in these statistically significant differences (See Appendix I). Bonferroni correction was used to control Type I error when reading the results of pairwise comparisons in each IV. For factor 1, *Gender* ($F(1, 220) = 5.35$, $p < .05$) and *Ethnicity* ($F(9, 220) = 2.65$, $p < .01$) have contributed to the statistically significant differences in student attitudes. Table 24 shows that the mean score of the male participants in factor 1 was .14 lower than the female participants, and this difference was statistically significant with $p < .05$.

Table 24

Pairwise Comparisons of Gender for Factor 1

					95% CI	
DV	(I) Gender	(J) Gender	ΔM (I-J)	p^b	LL	UL
Factor 1	Male	Female	-.14	.022*	-.26	-.02

Note. DV = dependent variable. ΔM = mean difference. I-J = male mean score subtracts female mean score. CI = confidence interval. LL = lower level. UL = upper level.

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Table 25 shows the pairwise comparisons among the White, Chinese, and the other ethnical groups. Only the mean difference between the White participants and Chinese participants was statistically significant at the Bonferroni adjusted significance level of .017. It

indicates that the white participants on average rated this factor .28 higher than the Chinese participants with $p < .001$.

Table 25

Pairwise Comparisons among Ethnical Groups in Factor 1

Dependent Variable	(I) Ethnicity	(J) Ethnicity	ΔM (I-J)	Sig.
Factor1	White	Chinese	0.28	.000*
		The Others	0.08	1.000

Note. Based on observed means. ΔM = mean difference between I and J. Bonferroni correction was used to control Type I error. The new adjusted significance level is .017. *. The mean difference is significant at the .017 level.

For Factor 3, *Frequencies of Using Computers for Learning per Week* has caused this statistically significant difference with $F(4, 220) = 4.54$, $p < .01$. Table 26 shows that the participants who never used computers for learning on average rated this factor 1.27 lower than the participants who spent more than 10 hours by using computers for learning per week. This mean difference was statistically significant with $p < .0125$.

Table 26

Pairwise Comparisons of Frequencies of Using Computers for Learning per Week for Factor 3

					95% CI	
DV	(I)	(J)	ΔM (I-J)	p^b	LL	UL
Factor 3	**Never**	1-3 hours per week	-1.10	0.026	-2.12	-0.08
		4-6 hours per week	-1.13	0.013	-2.12	-0.14
		7-9 hours per week	-1.09	0.020	-2.08	-0.10
		10 + hours per week	**-1.266**	**0.003***	-2.25	-0.28

Note. DV = dependent variable. ΔM = mean difference. I-J = mean score of I subtracts mean score of J. CI = confidence interval. LL = lower level. UL = upper level.
*. The mean difference is significant at the .0125 level.
b. Adjustment for multiple comparisons: Bonferroni.

For Factor 4, *Ethnicity* ($F(9, 269) = 3.59$, $p < .0056$) and *Frequencies of Playing Online/Video/Mobile Games per Week* ($F(4, 269) = 3.23$, $p = .013$) have contributed to the statistically significant difference among the participants. Table 27 shows that the white participants on average rated this factor .37 higher than the Chinese participants. This mean difference was statistically significant with $p < .001$ at the Bonferroni corrected alpha level of .017. The combination of all other ethnical groups also on average rated this factor .35 higher than the Chinese participants with $p < .01$. The statistically significant group differences suggested that Chinese participants found the least pleasure in learning experiences in this factor compared to the other ethnical groups of participants.

Table 27

Pairwise Comparisons of Ethnicity for Factor 4

Dependent Variable	(I) Ethnicity	(J) Ethnicity	ΔM (I-J)	Sig.
Factor4	Chinese	White	-0.37	.000*
		The Others	-0.35	.003*

Note. Based on observed means. ΔM = mean difference between I and J. Bonferroni correction was used to control Type I error. The new adjusted significance level is .017. *. The mean difference is significant at the .017 level.

Table 28 shows the results of pairwise comparisons between participants who spent 1-5 hours in playing online/video/mobile games per week and the other groups of participants in this IV. The reason for choosing this group of participants as the reference group was that the majority of participants in this research fell into this category. Table 28 indicates that the participants who never played the online/video/mobile games on average rated this factor .36 higher than the participants who played games for 1-5 hours per week. This mean difference was statistically significant with $p < .0125$ at the Bonferroni adjusted alpha level.

Table 28

Pairwise Comparisons of Frequencies of Playing Online/Video/Mobile Games per Week for Factor 4

					95% CI	
DV	(I)	(J)	ΔM (I-J)	p^b	LL	UL
Factor 4	**1-5 hours**	6-10 hours	0.00	1.000	-0.25	0.24
		11-15 hours	0.18	1.000	-0.34	0.71
		16 + hours	-0.08	1.000	-0.49	0.32
		Never	**-0.36**	**0.010***	-0.66	-0.05

Notes. DV = dependent variable. ΔM = mean difference. I-J = mean score of I subtracts mean score of J. CI = confidence interval. LL = lower level. UL = upper level.
*. The mean difference is significant at the .0125 level.
b. Adjustment for multiple comparisons: Bonferroni.

Comparisons between Chinese International and U.S. Domestic Undergraduates in the Modified Survey

In order to answer the sub-question under Research Question 2, one-way ANOVA was conducted between 82 Chinese international undergraduates and 119 U.S. domestic undergraduates from the sample. These 119 U.S. domestic undergraduates were selected from the sample with similar gender and age distributions as the Chinese international undergraduates had. The descriptive statistics of these two groups of participants are shown in Table 29 – 31.

Table 29

Demographic Information of the Chinese International Undergraduates and U.S. Domestic Undergraduates

	Age				Gender	
	18-19	20-21	22-23	24+	Male	Female
Chinese (N = 81)	27	25	18	11	43	38
%	33.3	30.9	22.2	13.6	53.1	46.9
US (N = 119)	30	53	21	15	74	45
%	25.2	44.5	17.6	12.6	62.2	37.8

Table 30

Academic Information of the Chinese International Undergraduates and U.S. Domestic Undergraduates

	Academic Year				
	1st	2nd	3rd	4th	5th +
Chinese (N = 81)	22	20	16	13	10
%	27.2	24.7	19.8	16.0	12.3
US (N = 119)	21	23	43	22	10
%	17.6	19.3	36.1	18.5	8.4

Major	Chinese	%	United States	%
Architecture	2	2.5	5	4.2
Arts	3	3.7	10	8.4
Business	30	37.0	16	13.4
Education	7	8.6	7	5.9
Engineering	21	25.9	8	6.7
Health Professions	1	1.2	10	8.4
Health, Sport, and Exercise Sciences			10	8.4
Humanities & International Studies	1	1.2	4	3.4
Journalism & Mass Communications	1	1.2	5	4.2
Law (Pre-Law)			3	2.5
Medicine (Pre-Med)	2	2.5	4	3.4
Music	2	2.5	1	0.8
Natural Science & Math	3	3.7	5	4.2
Nursing			2	1.7
Pharmacy	1	1.2		
Public Affairs & Administration			1	0.8
Social & Behavioral Sciences	2	2.5	26	21.8
Social Welfare	1	1.2	1	0.8
Undecided	4	4.9	1	0.8
Total	81	100.0	119	100.0

Table 31

Gamification Background of the Chinese International Undergraduates and U.S. Domestic Undergraduates

	Frequency of Using Computer for Learning per Week			
	1-3 hours/week	4-6 hours/week	7-9 hours/week	10+ hours/week
Chinese (N = 81)	5	16	17	43
%	6.2	19.8	21.0	53.10
US (N = 119)	11	38	32	37
%	9.2	31.9	26.9	31.1

	Gamification Experiences			
	Massive Multiplayer Online Games	Massive Open Online Courses	Online Shopping	Social Network Sites
Chinese (N = 81)	38	25	72	58
%	46.9	30.9	88.9	71.6

US (N = 119)	66	45	110	114
%	55.5	37.8	92.4	95.8

	Years of Playing Video/Online/Mobile Games			
	1-3 years	4-6 years	7+ years	Never
Chinese (N = 81)	20	18	33	10
%	24.7	22.2	40.7	12.3
US (N = 119)	11	15	69	24
%	9.2	12.6	58.0	20.2

	Frequency of Playing Video/Online/Mobile Games per Week				
	1-5 hours	6-10 hours	11-15 hours	16+ hours	Never
Chinese (N = 81)	35	16	6	11	13
%	43.2	19.8	7.4	13.6	16.0
US (N = 119)	59	20	2	2	36
%	49.6	16.8	1.7	1.7	30.3

	Gamer Type					
	Story-driven Solo Gamer	Social Gamer	Solo-limited Gamer	Hardcore Online Gamer	Control/identity Solo Gamer	Casual Gamer
Chinese (N = 81)	14	12	7	6	15	17
%	17.3	14.8	8.6	7.4	18.5	21.0
US (N = 119)	10	29	11	4	13	29
%	8.4	24.4	9.2	3.4	10.9	24.4

Note. 10 (12.3%) Chinese international participants indicated that they were not belong to any gamer type listed above. 23 (19.3%) U.S. domestic students also considered themselves not belong to any gamer type listed above.

The mean score of Chinese international undergraduates for each factor was consistently lower than the U.S. domestic undergraduates, and these mean differences were statistically significant at the significance level of .05 (See Table 32). In factor 1, $M_{Chinese} = 3.33$, $M_{US} = 3.60$, $F(1,197) = 21.00$, $p < .001$; in factor 2, $M_{Chinese} = 3.15$, $M_{US} = 3.37$, $F(1, 198) = 8.36$, $p < .01$; in factor 3, $M_{Chinese} = 3.11$, $M_{US} = 3.17$, $F(1, 198) = 4.48$, $p < .05$; and in factor 4, $M_{Chinese} = 2.96$, $M_{US} = 3.35$, $F(1, 198) = 34.25$, $p < .001$.

Table 32

Outputs of One-Way ANOVA between Chinese International Undergraduates and U.S. Domestic Undergraduates

DV	Group	N	M	SD	F(1, 198)	*p*
Factor 1	US Domestic	119	3.60	0.36	21.00	.000***
	Chinese International	80	3.33	0.47		
	Total	199	3.49	0.43		
Factor 2	US Domestic	119	3.37	0.55	8.36	.004**
	Chinese International	81	3.15	0.52		
	Total	200	3.28	0.55		
Factor 3	US Domestic	119	3.17	0.44	4.48	.035*
	Chinese International	81	3.04	0.41		
	Total	200	3.11	0.43		
Factor 4	US Domestic	119	3.35	0.42	34.25	.000***
	Chinese International	81	2.96	0.51		
	Total	200	3.19	0.49		

Note. DV = dependent variable. M = mean. SD = standard deviation. The significance level was .05. *p < .05. ** p < .01. *** p < .001.

Discussions

This chapter discusses the findings that were reported in Results section, such as the overall scale, the usefulness of this instrument, and the implications of the comparisons between the Chinese international undergraduates and the U.S. domestic participants in this research. This chapter also discusses whether participants' previous gamification experiences would affect participants' attitudes towards pleasurability in learning. Additionally, this chapter compares the findings in Results section with the existing literature on educational gamification.

Measurement Model

The purpose of designing gamified learning environments is to engage students in learning. But what game factors are the most effective to engage the learners are not clear from current literature given that most research was case studies and could not be generalized to different learners (Hamari et al., 2014; Nah et al, 2014). The usefulness of this instrument is that it could tell the instructional designers and instructors what game factors are highly valued by their target learners before they design the gamified learning environments.

Students often learn by playing in their daily lives, but they need to adopt a different set of learning methods in formal learning environments. Is this difference in the learning methods necessary? Gamification is a possible approach to narrow the gap between these two different sets of learning methods. As game factors are meant to provide pleasurable experiences for the learners, investigating undergraduates' pleasurable learning experiences would be able to match these experiences with the game factors. Nokia research team has used PLEX in the development of game applications (Arrasvuori et al., 2011). 22 categories of PLEX have been selected as the general fun and playful user experiences. After the interviews and the focus groups, 16 categories in PLEX were selected to be suitable in learning environments and 2 new categories

Visualization and *Interaction* were added to the list. The 18 categories of pleasurable learning experiences (PLLEXs) used in this research have concrete descriptions of learning activities that can be matched with the definition of each PLLEX category. These concrete PLLEX categories are different from the categories in PLEX, which are abstract feelings.

The EFA results suggest that college undergraduate students' pleasurability in learning could be analyzed from 4 aspects: *Preferences for Instructions*, *Preferences for Teaching Styles*, *Preferences for Activities*, and *Preferences for Learning Effectiveness*. Comparing with the 18 categories in PLLEX, these 4 aspects have explained these pleasurable feelings:

Factor 1: *Preferences for Instructions*

PLLEX	Definition	Survey Item
Control	Dominating, commanding, regulating	12. I like being able to check my progress. 20. I like instructors providing clear syllabus to us. 19. I like instructors to provide clear rubrics to us. 5. I like instructors to share class materials with us. 21. I like learning knowledge and skills that are related to the final exam of the course.
Submission	Being part of a larger structure	9. I like instructors to give us guidance/directions when we are conducting in-class activities. 10. I like receiving instructors' comments on my assignments. 13. I like getting timely feedbacks on my performances from the instructors.

Note. Item 13 and item 10 also belong to Control category. Survey items are listed by their factor loadings. The higher the item is, the more relevant this item is associated with the factor.

Factor 2: *Preferences for Teaching Styles*

PLLEX	Definition	Survey Item
Humor	Fun, joy, amusement, jokes,	28. I like instructors to make

	gags	jokes during instructions. 27. I like instructors to use funny pictures or quotations during instructions.
Sensation	Excitement by stimulating senses	29. I like instructors using videos during instructions.
Visualization	Recalling or forming mental images	

Note. Survey items are listed by their factor loadings. The higher the item is, the more relevant this item is associated with the factor.

Factor 3: *Preferences for Activities*

PLLEX	Definition	Survey Item
Interaction	An reciprocal action, effect, or influence	18. I like interactive in-class activities. 4. I like instructors to use interactive technology (e.g., i-clickers and social media tools).
Fellowship	Friendship, communality or intimacy	15. I like collaborating with my classmates. 30. I like building personal connections with my classmates.
Discovery	Finding something new or unknown	7. I like trying something new.
Challenge	Testing abilities in a demanding task	8. I like challenges.

Note. Survey item 15 and survey item 30 also belong to Interaction category. Survey items are listed by their factor loadings. The higher the item is, the more relevant this item is associated with the factor.

Factor 4: *Preferences for Learning Effectiveness*

PLLEX	Definition	Survey Item
Fellowship	Friendship, communality or intimacy	23. I like building personal connections with instructors. 22. I like instructors' attentions in class.
Control	Dominating, commanding, regulating	14. I like reviewing contents. 17. I like being self-motivated. 24. I like instructors being accessible.

Note. Survey item 24 also belongs to Fellowship categor*y*. Survey items are listed by their factor loadings. The higher the item is, the more relevant this item is associated with the factor.

By understanding the relationships between the survey items and PLLEX categories, the abstract pleasurable feelings become concrete learning activities that could be accomplished by designing gamified instructional environments for the learners. Based on the survey items and the definitions of game factors, the divisions of the game factors could be concluded as the match-up in Table 33.

Table 33

Matrix of 4 Subscales from EFA and 7 Game Factors

PLLEX Categories	EFA Subscales	Game Factors
Control Submission	Preferences for Instructions	Feedback Points Progress Storytelling (rules)
Humor Sensation Visualization	Preferences for Teaching Styles	Storytelling (presentation)
Challenge Discovery Fellowship Interaction	Preference for Activities	Badges Levels
Control Fellowship	Preferences for Learning Effectiveness	Storytelling (supportive information)

The game factor *Storytelling* could be matched with three subscales, but each *Storytelling* has a different focus. *Storytelling* in the subscale *Preferences for Instructions* emphasizes the *rules* of the gamified learning environments, such as syllabus of the course, rubrics for the assignments, and directions of tasks; *Storytelling* in the subscale *Preferences for Instructors' Teaching Style* focuses on the ways that instructors *present* the contents. For example, the instructors could use more multimedia resources to present their instructional materials; *Storytelling* in the subscale *Preferences for Learning Effectiveness* puts emphasis on designing *scaffolding materials* for the learners, such as providing background knowledge of the new contents and instructors' virtual office hours with the learners.

The Usefulness of this Instrument

The rankings of the mean score of each subscale can be interpreted as students' preferences for the game factors that should be included in gamified learning environments. The higher the mean score is, the more emphases should be given to the game factors that are matched with the subscale(s). Using this research as an example (See Table 34), the mean score of each subscale is $M_{Factor1} = 3.53$, $M_{Factor2} = 3.30$, $M_{Factor3} = 3.09$, and $M_{Factor4} = 3.22$, which indicates that the participants in this research had higher demands for the design of *Feedback*, *Points*, *Progress*, and *Storytelling (rules)* in gamified learning environments. But the design of *Badges* and *Levels* are less important for this particular group of participants as this subscale has the lowest mean score.

Table 34

Descriptive Statistics of Subscales of the Finalized Survey

DV	N	M	SD	Skewness	Kurtosis
Factor 1	278	3.53	0.42	-0.91	0.83
Factor 2	279	3.31	0.58	-0.71	0.80
Factor 3	279	3.09	0.47	-0.23	-0.18
Factor 4	279	3.22	0.47	-0.42	0.96
Total	278				

Note. N = number. M = mean. SD = standard deviation.

However, the findings of this research also suggest that when splitting the participants by their different racial backgrounds, different groups of participants had different focuses on the design of game factors. Based on the results of MANOVA, the group differences are statistically significant. White participants rated these 4 subscales almost the same pattern as the overall population did, but black participants rated the four subscales as $M_{Factor1} = 3.29$, $M_{Factor2} = 3.16$, $M_{Factor3} = 3.27$, and $M_{Factor4} = 3.16$ (See Table 35). The mean scores indicated by this group of students suggest that when designing the gamified learning environments for them, besides the

game factors related to Factor 1, the instructional designers and instructors also need to focus on the design of *Badges* and *Levels*.

Table 35

Descriptive Statistics of Black or African American Participants

DV	N	M	SD	Skewness	Kurtosis
Factor 1	9	3.29	0.58	-1.42	2.66
Factor 2	9	3.16	0.65	-0.21	-0.26
Factor 3	9	3.27	0.67	-0.98	-0.19
Factor 4	9	3.16	0.43	0.37	-0.50
Total	9				

Note. N = number. M = mean. SD = standard deviation.

In addition, Chinese participants rated these four subscales differently from white and black participants with $M_{Factor1} = 3.33$, $M_{Factor2} = 3.15$, $M_{Factor3} = 3.04$, and $M_{Factor4} = 2.96$ (See Table 36). This result indicates that the design of *Storytelling (supportive information)* might not engage these Chinese participants as much as the good design of *Feedback*, *Points*, *Progress* and *Storytelling (rules)* in gamified learning environments.

Table 36

Descriptive Statistics of Chinese Participants

DV	N	M	SD	Skewness	Kurtosis
Factor 1	81	3.33	0.47	-0.52	0.38
Factor 2	82	3.15	0.52	-0.46	0.60
Factor 3	82	3.04	0.40	0.06	0.35
Factor 4	82	2.96	0.50	-0.41	1.92
Total	81				

Note. N = number. M = mean. SD = standard deviation.

From the examples given above, this survey could help the instructional designers and instructors specifically know the critical game factors they should emphasize for different groups of learners when they design gamified learning environments. The findings of this research suggest that there is no universal design feature for different learners. The 4 subscales in this

survey can help the instructional designers and instructors know about their learners' preferences for the design of gamified learning environments. By understanding target learners' different preferences for PLLEXs, the instructional designers and instructors can become more sensitive when they have learners who are from diverse backgrounds.

Different Preferences between Chinese International Undergraduates and U.S. Domestic Undergraduates

The results of one-way ANOVA between Chinese international undergraduates and U.S. domestic undergraduates showed that Chinese international undergraduates tended to feel less pleasure in learning compared to the U.S. domestic undergraduates (See Table 30). Both of the Chinese and U.S. undergraduate students agreed the highest with the pleasure they felt in their *Preferences for Instructions*, but Chinese international undergraduates did not find as much pleasure as the U.S. domestic undergraduates did in their *Preferences for Learning Effectiveness*. This difference suggests that the Chinese international undergraduates are less likely to build connections with the instructors or do not have strong motivations to study by themselves after class, compared with domestic students in the United States. Furthermore, this difference also implies that the supportive information provided by the instructional designers and instructors in gamified learning environments may not be effective to engage Chinese international undergraduates.

This difference might be caused by the language barriers of this group of participants as the descriptive statistics showe that most Chinese international undergraduates participated in this research are freshmen or sophomore at KU. It could also because of the cultural difference as in Chinese cultures, students tend to show great respects to the professors and regard them as superior figures. Therefore, they would be shy to build connections with the instructors.

By understanding the different preferences for PLLEX between the Chinese international students and the U.S. domestic students, one message that delivers to the instructional designers and instructors is that the design of game factors should not stay the same for different learners. The successful instructional experiences with previous learners would not necessarily be effective with the new group of learners. As more and more international students have been enrolled into the colleges and universities in the U.S., the sensitiveness of different learning preferences among students who come from different countries of origins should be strengthened among the instructional designers and instructors.

Contributions to the Theoretical Frameworks in Educational Gamification

This research finds that aside from the statistically significant differences in *Ethnicity*, participants also differ significantly in their attitudes towards pleasure in learning based on their *Gender* and *Frequencies of Using Computers for Learning Purposes*. The findings of *Gender* difference in this research are consistent with the current literature on gamification, which indicates that females tend to have better experiences with gamification than males (Hamari et al., 2014). This difference is found as statistically significant in participants' *Preferences for Instructions* and *Preferences for Instructors' Teaching Styles*. This difference suggests that females are more sensitive to good design of *Feedback*, *Points*, *Progress*, and *Storytelling (rules and presentations)* than males.

Ethnicity and *Frequencies of Using Computers for Learning Purposes per Week* are two new IVs that have not been investigated in current literature on educational gamification. The statistically significant differences that are caused by *Ethnicity* could make future instructional designers and instructors become aware of the different preferences for game factors in the groups of students who come from different racial backgrounds. As students' *Frequencies of*

Using Computers for Learning Purposes per Week yields statistically significant differences in the *Preferences for Activities* subscale, it indicates that the more often students use their computers for learning, the more likely they would like *Badges* and *Levels* in gamified learning environments.

Gamer Types was mentioned in Landers & Callan (2011) as different gamers might have different starting point in their attitudes towards gamification. However, this study finds that *Gamer Types* does not affect participants' preferences for the design of game factors. This finding suggests that students who have played online/video/mobile games in the past or not could equally find pleasure in gamified learning environments.

Age also does not result in any statistically significant difference in participants' attitudes in this research. Koivistro & Hamari (2014) found that the ease of use in gamification would decline with the growth of age. But in this research, 279 participants' attitudes towards game factors are similar. This might be due to the fact that the age difference of students who participated in this research was only 6 years. If comparing the participants with wider age ranges, there might be some statistically significant differences.

Summary

The purpose of this research is to develop an instrument that can help instructional designers and instructors to decide the critical game factors in their gamified learning environments for college undergraduate students. Studying educational gamification from college students' perspective is important because the purpose of designing gamified learning environments is to engage the students and thus, to let them experience pleasure in meaningful learning.

The mystery of education is that many instructors believe that they know the best about their students, so they insist on using the same methods when they teach students. But it is inevitable that technology advances have changed the ways that today's college students learn. Students could Google on their laptops or tablets in classes when the instructors talk about new terminologies and select the best answers by looking at the numbers of "thumbs-up" (agreements by other users). Students would also "tweet" on social media to share educational videos and news with their classmates and instructors. Can the old-school teaching methods still engage today's college students?

There is always a huge gap between how students learn in their daily lives and how they learn in the academic learning environments. Is this learning curve necessary? Play is in the nature of people, and this is the natural way that people learn new things. Well-designed gamified learning environments would bring up more pleasurability to the students. The literature has reiterated the importance of student engagement in learning. By using this instrument, the instructors would be able to better understand their students' needs for pleasurability in learning and therefore, they could better engage the students.

After interviewing and conducting the focus groups with the KU undergraduate students, 36 common pleasurable learning experiences were generated. Initially, 18 PLLEX categories were used to map these pleasurable learning experiences. In the finalized scale, only 22 survey items remained and could represent 9 PLLEX categories, which suggest that these 9 PLLEX categories are the major pleasurable feelings that undergraduate students experience in their academic learning environments.

PLLEX is adapted from PLEX (Arrasvuori et al., 2011). Nokia research team used PLEX as a design framework in developing their game applications. A Korean researcher also used this

framework to design his gamified class at a university in South Korea (Kim, 2013). Well-designed educational gamification should bring PLLEXs to learners with the purpose of promoting student engagement. Given that the definition of gamification is to use game factors in non-game environments, selecting appropriate game factors for different learning contexts is very important in the design of educational gamification because different game factors would stimulate different PLLEXs. Using PLEX as a design framework would help this research generate more comprehensive PLLEXs to map with the game factors.

A 4-Point Likert agree/disagree scale was created to investigate college undergraduate preferences for pleasurable learning experiences. Factor analysis has revealed that 4 factors could best describe the variance in the scale. The items under each factor were mapped with several game factors. The instructional designers and instructors could distribute this scale to their students at the beginning of the classes, and then calculate the mean score of each subscale. The ranking of the mean scores can tell the instructors that the subscales of pleasurable learning experiences the students highly agree with, which can also tell the instructors the game factors that they need to emphasize in designing gamified learning environments for their students.

The differences between the Chinese international undergraduates and the U.S. domestic undergraduates indicate that the preferences for each subscale might differ among the students who are from different countries of origins. This difference could result from their different learning purposes or different cultural traditions and values. Further research is warranted to compare Chinese college students who study in China and Chinese college students who study abroad.

The limitation of this study is that because of the limited time and sources, only 279 undergraduates were recruited and not enough sample size to conduct confirmatory factor

analysis (CFA) to test whether the model fits different samples, such as Freshmen vs. Senior, Humanities vs. STEM, White vs. Hispanic. Another limitation is that only 3 survey items were under Factor 2, more items that are related to instructors' teaching styles should be created to represent that subscale. Future research could be recruiting more college students and use CFA to check the measurement model fit. Invariance tests could also be conducted if the new sample involves students from different countries of origins.

References

Andrade, M. S. (2006). International students in English-speaking universities Adjustment factors. *Journal of Research in International Education*, *5*(2), 131-154.

Arrasvuori, J., Boberg, M., Holopainen, J., Korhonen, H., Lucero, A., & Montola, M. (2011, June). Applying the PLEX framework in designing for playfulness. In*Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces* (p. 24). ACM.

Barber, W., & Badre, A. (1998, June). Culturability: The merging of culture and usability. In *Proceedings of the 4th Conference on Human Factors and the Web* (pp. 1-14).

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical assessment, research & evaluation*, *18*(6), 1-13.

Biro, G. I. (2013). Ready, Study, Share: An Inquiry Into the Didactic Approach of Gamification With a Special View to the Possible Application in Higher Education, 1st Annual International Interdisciplinary Conference, Conference Proceedings, Vol.2, EGALITE, European Scientific Institute Publishing.

Bonk, C. J., & Cunningham, D. J. (1998). Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools. *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse*, *25*.Campbell, D. T. & Stanley, J.C. (1963). Experimental and quasi-experimental designs for research. Boston: Houghton Mifflin.

Carmines, E. G. & Zeller, R.A. (1991). *Reliability and validity assessment*. Newbury Park: Sage Publications.

Chen, H. L., Lattuca, L. R., & Hamilton, E. R. (2008). Conceptualizing engagement: Contributions of faculty to student engagement in engineering.*Journal of Engineering Education*, *97*(3), 339-353.Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: Harper and Row.

de Araujo, A. A. (2011). Adjustment issues of international students enrolled in American colleges and universities: A review of the literature. *Higher Education Studies*, *1*(1), p2.DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research*, *3*(1), 60-95.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15). ACM.

DeVellis, Robert F. (2003). *Scale development: Theory and applications* (Vol. 26). Sage publications.

Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, *63*, 380-392.

Erenli, K. (2013). The Impact of Gamification-Recommending Education Scenarios. *International Journal of Emerging Technologies in Learning (iJET)*, *8*(2013), 15-21.

Gay, G. & Hembrooke, H. (2004). *Activity-centered design: An ecological approach to designing smart tools and usable systems*. MIT Press.

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.

Glover, I. (2013). Play as you learn: gamification as a technique for motivating learners.

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). Does Gamification Work?--A Literature Review of Empirical Studies on Gamification. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on* (pp. 3025-3034). IEEE.

Higher Education Funding Council of England. (2005). HEFCE strategy for e-learning. Retrieved January 16, 2006 from http://www.hefce.ac.uk/pubs/hefce/2005/05_12/

Hofstede, G., & Hofstede, G. J. Cultures and Organizations. Software of the Mind. 2005. *McGrawHill, New York*.

Horstmanshof, L and Zimitat, C, Future time orientation predicts academic engagement among first year university students, British Journal of Educational Psychology, 77, (3) pp. 703-718. ISSN 0007-0998 (2007)

Huitt, William G. (1998). *Internal and External Validity*. http://www.valdosta.peachnet.edu/~whuitt/psy702/intro/valdgn.html

Jensen, M. (2012). Engaging the learner: gamification strives to keep the user's interest. *T+D*, *66*(1), p40-44.

Jordan, P. W. Designing pleasurable products: an introduction to the new human factors. 2000.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Kim, S. (2013). Analysis of Engineering Students' Needs for Gamification based on PLEX Model. *Journal on Knowledge and Data Engineering (JKDE)*, *1*(1), 1–7.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, *35*, 179-188.

Korhonen, H., Arrasvuori, J., & Väänänen-Vainio-Mattila, K. (2010, December). Analysing user experience of personal mobile products through contextual factors. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia* (p. 11). ACM.

Krause, K. (2005). Understanding and promoting student engagement in university learning communities. *Paper presented as keynote address: Engaged, Inert or Otherwise Occupied*.

Krause, K. L., & Coates, H. (2008). Students' engagement in first-year university. *Assessment & Evaluation in Higher Education*, *33*(5), 493-505.

Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. *Journal of College Student Development*, *50*(6), 683-706.

Landers, R. N., & Callan, R. C. (2011). Casual social games as serious games: The psychology of gamification in undergraduate education and employee training. In *Serious games and edutainment applications* (pp. 399-423). Springer London.

Lee, J. J. (2010). International students' experiences and attitudes at a US host institution: Self-reports and future recommendations. *Journal of Research in International Education*, *9*(1), 66-84.

Lindholm, C., & Keinonen, T. (2003). *Mobile Usability: How Nokia Changed the Face of the Cellular Phone*. McGraw-Hill, Inc..

Maxwell, Joseph A. (2013). Qualitative research design: An interactive approach (3rd edition). Thousand Oaks, CA: Sage Publications.

McGonigal, J. (2011). Reality is broken: Why games make us better and how they can change the world (1st ed.). New York, NY, USA: The Penguin Press.

Nah, F. F. H., Zeng, Q., Telaprolu, V. R., Ayyappa, A. P., & Eschenbrenner, B. (2014). Gamification of Education: A Review of Literature. In *HCI in Business: First International Conference, HCIB 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings* (pp. 401-409). Springer International Publishing.

Hoft, N. L. (1995). *International technical communication: How to export information about high technology*. John Wiley & Sons, Inc..

Payette, N., & Hardy-Vallée, B. (2008). *Beyond the brain: Embodied, situated and distributed cognition*. Cambridge Scholars Pub.

Pedreira, O., García, F., Brisaboa, N., & Piattini, M. (2015). Gamification in software engineering–A systematic mapping. *Information and Software Technology*, *57*, 157-168.

Polit, D. F., Beck, C. T. (2012). *Nursing Research: Generating and Assessing Evidence for Nursing Practice, 9th ed*. Philadelphia, USA: Wolters Klower Health, Lippincott Williams & Wilkins.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.*, *57*, 87-115.

Sleeter, Christine E. & Grant, Carl A. (2009, 6th Edition). Making choices for multicultural education: Five approaches to race, class, and gender. New York, NY: John Wiley and Sons, Inc.

Stevens, J. P. (1992). Applied multivariate statistics for the social sciences (2nd edition). Hillsdale, NJ: Erlbaum.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, *2*, 53.

Tiger, L. (1992). *The pursuit of pleasure*. Transaction Publishers.

Urh, M., Vukovic, G., Jereb, E., & Pintar, R. (2015). The Model for Introduction of Gamification into E-learning in Higher Education. *Procedia - Social and Behavioral Sciences*, *197*(February), 388–397. http://doi.org/10.1016/j.sbspro.2015.07.154

Xiang, B., & Shen, W. (2009). International student migration and social stratification in China. *International Journal of Educational Development*, *29*(5), 513-522.

Xu, Y. (2011). Literature review on web application gamification and analytics. *Honolulu, HI*, 11-05.

Yamagata-Lynch, L. C. (2010). *Activity systems analysis methods: Understanding complex learning environments*. Springer Science & Business Media.

Zhai, L. (2004) Studying International Students: Adjustment Issues and Social Support. *depression*, *11*(1).

Appendices

Appendix A

Student Interview Protocol

Goal: Understand target audience holistically

Objectives: Find out students who are from different majors in-class activities, assignment requirements, and online/blended learning experiences.

For undergraduates:

Q1: Please introduce yourself (age, major, ethnicity background, current studying level).

Q2: What are the most frequent in-class activities the instructors conduct in the classroom?

- What activities do you like/find the most useful? And why? (feelings, events, rewards, etc.)
- What activities do you dislike/find the least useful? And why? (feelings, events, punishments,etc.)

Q3: What assignment formats do your instructors use?

- What assignment format do you like/dislike? And why?

Q4: Do your instructors use Bb or other online learning management system to organize course materials?

- What supportive materials do you find the most/least useful? And why?

Q5: Do your instructors use technology, such as PowerPoints, streaming videos, and other multimedia resources, in their courses?

- What is your opinion for the effectiveness of their usage of technology?

Q6: What improvement would you like to see in the future for your classes?

For graduate teaching assistants (GTA):

Q1: Please introduce yourself (age, major, ethnicity background, current studying level).

Q2: What are the most frequent in-class activities do you conduct in the classroom?

- What activities do undergraduates like/find the most useful?
- What activities do undergraduates dislike/find the least useful?

Q3: What assignment formats do you use?

- What assignment format do you find most effective in assessing undergraduates?

Q4: Do you use Bb or other online learning management system to organize course materials?

- Do you think they are effective in presenting contents, communicating with undergraduates, assessing undergraduates?

Q5: Do you use technology, such as PowerPoints, streaming videos, and other multimedia resources, in your class?

- What is your opinion for the effectiveness of technology?

Appendix B

Questions for Focus Groups

Goal: Understand target audience holistically

Objectives: Find out students' shared pleasurable learning experiences in both online and face-to-face instructions from who are from different majors.

Tools: Post-it, maker pens, whiteboard, PLEX cards (cards with 22 categories of fun and pleasure factors)

Process: (1) Students will use Post-it to stick their answers to the whiteboard; (2) students will refer to some cards printed with pleasurable factors to brainstorm their solutions for situated learning environments.

Q1: Please introduce yourself (age, major, ethnicity background, current studying level).

Q2: What kind of learning experiences do you consider as pleasurable in your face-to-face classes? And why? (Post-it)

Q3: What kind of learning experiences do you dislike in your face-to-face classes? And why? (Post-it)

Q4: What kind of learning experiences do you consider as pleasurable in your online classes? And why? (Post-it)

Q5: What kind of learning experiences do you dislike in your online classes? And why? (Post-it)

Q6: Imagine you are an instructor in a face-to-face philosophy undergraduate class, how would you increase students' pleasurable learning experiences? (PLEX cards)

Q7: Imagine you are an instructor in an online research methods undergraduate class, how would you increase students' pleasurable learning experiences? (PLEX cards)

Appendix C

Normality Tests of 34 Survey Items in the Pilot Test

Survey Item	Skewness (N = 88)	Kurtosis (N = 88)
1. I like taking notes during instructions.	-0.63	0.18
2. I like instructors to use PowerPoint slides.	-0.74	1.75
3. I like instructors to write the contents on the white/black board.	-0.51	-0.14
4. I like learning new contents outside of the class while doing activities in class.	-0.52	0.44
5. I like instructors to use interactive technology (e.g., i-clickers and social media tools).	-0.50	0.16
6. I like instructors to share class materials.	-0.88	1.28
7. I like previewing new content.	-0.30	-0.30
8. I like trying new things.	-0.40	-0.71
9. I like challenges.	-0.16	-0.60
10. I like instructors to give us guidance/directions when we are doing in-class activities.	-0.49	-0.75
11. I like receiving instructors' comments on my assignments.	-0.38	-1.90
12. I like participating in virtual discussion groups, e.g., discussion board on Blackboard application.	0.18	-0.50
13. I like being able to check my progress.	-1.26	0.65
14. I prefer subjective assessments, such as essays and short answers.	0.14	-0.78
15. I prefer objective assessments, such as multiple-choice questions and tests.	-0.62	0.02
16. I like getting quick feedback on my performances from the instructors.	-0.67	-0.56
17. I like reviewing contents.	-0.48	0.19
18. I like collaborating with my classmates.	-0.42	-0.24
19. I like learning new materials that are related to my prior knowledge.	-0.04	-1.07
20. I like self-paced learning experiences.	-0.36	0.74
21. I am self-motivated.	-0.51	-0.84
22. I like interactive in-class activities.	-0.66	0.78

23. I like learning knowledge and skills that can apply directly to the real world problems.	-0.93	-0.10
24. I like instructors to provide clear rubrics to us.	-0.62	-0.77
25. I like instructors to provide clear syllabus to us.	-0.93	-0.23
26. I like learning knowledge and skills that are related to the final exam of the course.	-1.24	1.63
27. I like instructors' attentions.	-0.46	0.24
28. I like building personal connections with instructors.	-0.17	-0.56
29. I like instructors being accessible.	-0.24	-1.41
30. I like knowing other students' grades on the assignments.	0.16	-0.76
31. I like the feeling of completing a major task.	-0.87	-0.25
32. I like instructors to use scenarios during instructions.	-0.78	0.82
33. I like instructors to use funny pictures and relevant quotations during instructions.	-0.99	0.62
34. I like instructors to make jokes and tell stories during instruction.	-1.40	1.96

Appendix D

Factor Correlation Matrix in the Pilot Test

Factor	1	2	3	4
1	1.00	0.38	0.42	0.12
2	0.38	1.00	0.25	0.11
3	0.42	0.25	1.00	0.26
4	0.12	0.11	0.26	1.00

Notes. Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalization.

Appendix E

Cronbach's Alpha Scale Reliability Tests of the Substantial Scales in the Pilot Test

Substantial Scale	Cronbach's Alpha	N of Items
Factor 1	.84	7
Factor 2	.82	9
Factor 3	.76	7
Factor 4	.69	5

Appendix F

EFA Outputs of the Overall Scale in the Pilot Test

Tests of Between-Subjects Effects in the Pilot Test

Source	Dependent Variable	df	F	*p*.	η^2
Corrected Model	Factor1	45	1.54	.083	.628
	Factor2	45	1.08	.398	.543
	Factor3	45	1.60	.065	.637
	Factor4	45	1.47	.106	.618
Age	Factor1	4	0.58	.680	.053
	Factor2	4	0.82	.522	.074
	Factor3	4	2.54	.055	.198
	Factor4	4	0.64	.637	.059
Gender	Factor1	1	1.77	.191	.041
	Factor2	1	0.07	.799	.002
	Factor3	1	3.26	.078	.074
	Factor4	1	0.65	.424	.016
Ethnicity	Factor1	4	5.94	.001**	.367
	Factor2	4	1.16	.344	.101
	Factor3	4	1.36	.264	.117
	Factor4	4	3.03	.028*	.228
Current Study Year	Factor1	4	1.63	.185	.137
	Factor2	4	0.13	.972	.012
	Factor3	4	0.43	.786	.040
	Factor4	4	0.56	.695	.052
Major	Factor1	17	1.00	.482	.292
	Factor2	17	0.90	.584	.271
	Factor3	17	0.63	.852	.206
	Factor4	17	0.22	.999	.084
Frequency of Using	Factor1	3	3.38	.027*	.198
Computers for	Factor2	3	0.50	.683	.035
Learning per Week	Factor3	3	1.58	.209	.104
	Factor4	3	0.69	.563	.048
Years of Playing	Factor1	3	2.05	.121	.131
Online/Video/Mobile	Factor2	3	1.20	.323	.081
Games	Factor3	3	0.41	.750	.029
	Factor4	3	0.80	.499	.055
Frequency of Playing	Factor1	4	0.16	.955	.016
Online/Video/Mobile	Factor2	4	0.42	.792	.040

Games per Week	Factor3	4	1.37	.263	.118
	Factor4	4	0.38	.819	.036
Gamer Type	Factor1	5	1.28	.290	.135
	Factor2	5	0.47	.797	.054
	Factor3	5	0.94	.468	.102
	Factor4	5	1.48	.217	.153
Error	Factor1	41			
	Factor2	41			
	Factor3	41			
	Factor4	41			

Note. In corrected models, factor 1 R Squared = .628 (Adjusted R Squared = .219; factor 2 R Squared = .543 (Adjusted R Squared = .042); factor 3 R Squared = .637 (Adjusted R Squared = .239); factor 4 R Squared = .618 (Adjusted R Squared = .198). Using level of significance is $p < .05$. * $p < .05$, ** $p < .01$, *** $p < .001$.

Appendix G

Cronbach's Alpha Scale Reliability Test of the Overall Scale in the Modified Survey

Table G1

Cronbach's Alpha Reliability Statistics of the Overall Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.88	.89	30

Table G2

Survey Item-Total Statistics of the Overall Scale

Survey Item	Scale Mean if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 1	94.82	0.24	0.89
Item 2	94.66	0.30	0.88
Item 3	94.92	0.27	0.88
Item 4	95.01	0.27	0.88
Item 5	94.41	0.47	0.88
Item 6	94.76	0.45	0.88
Item 7	94.50	0.50	0.88
Item 8	94.59	0.39	0.88
Item 9	94.42	0.50	0.88
Item 10	94.32	0.53	0.88
Item 11	95.45	0.15	0.89
Item 12	94.28	0.47	0.88
Item 13	94.33	0.59	0.88
Item 14	94.71	0.47	0.88
Item 15	95.02	0.30	0.88
Item 16	94.48	0.51	0.88
Item 17	94.68	0.46	0.88
Item 18	94.86	0.30	0.88
Item 19	94.36	0.57	0.88
Item 20	94.28	0.53	0.88
Item 21	94.34	0.41	0.88
Item 22	94.69	0.52	0.88
Item 23	94.72	0.46	0.88
Item 24	94.42	0.57	0.88
Item 25	94.29	0.47	0.88

Item 26	94.54	0.60	0.88
Item 27	94.55	0.51	0.88
Item 28	94.54	0.42	0.88
Item 29	94.60	0.48	0.88
Item 30	94.67	0.47	0.88

Appendix H

Factor Loadings of Each Survey Item in the Modified Survey

Survey Item	Factor Loading			
	Factor 1	Factor 2	Factor 3	Factor 4
12. I like being able to check my progress.	0.83			
20. I like instructors to provide clear syllabus to us.	0.78			
19. I like instructors to provide clear rubrics to us.	0.74			
9. I like instructors to give us guidance/directions when we are doing in-class activities.	0.70			
5. I like instructors to share class materials.	0.67			
21. I like learning knowledge and skills that are related to the final exam of the course.	0.66			
10. I like receiving instructors' comments on my assignments.	0.65			
13. I like getting quick feedback on my performances from the instructors.	0.63			
28. I like instructors to make jokes and tell stories during instructions.		0.90		
27. I like instructors to use funny pictures and relevant quotations during instructions.		0.84		
29. I like instructors to use videos during instructions.		0.83		
26. I like instructors to use scenarios during instructions.		0.48		
18. I like interactive in-class activities.			0.79	
4. I like instructors to use interactive technology (e.g., i-clickers and social media tools).			0.75	
15. I like collaborating with my classmates.			0.65	
30. I like building personal connections with my classmates.			0.56	
7. I like trying new things.			0.56	
8. I like challenges.			0.50	
23. I like building personal connections with instructors.				0.82
22. I like instructors' attentions.				0.76
14. I like reviewing contents.				0.61
17. I am self-motivated.				0.57
24. I like instructors being accessible.				0.46

Appendix I

Main Effect of Each IV on Each DV in the Modified Survey

Source	DV	df	F	*p*	η^2
Corrected Model	Factor1	52	1.948***	.000	.315
	Factor2	52	1.368	.064	.244
	Factor3	52	1.872***	.001	.307
	Factor4	52	1.984***	.000	.319
Intercept	Factor1	1	652.030	.000	.748
	Factor2	1	317.127	.000	.590
	Factor3	1	354.864	.000	.617
	Factor4	1	454.737	.000	.674
Age	Factor1	3	1.618	.186	.022
	Factor2	3	.785	.503	.011
	Factor3	3	.485	.693	.007
	Factor4	3	.670	.571	.009
Gender	Factor1	1	5.349*	.022	.024
	Factor2	1	4.058*	.045	.018
	Factor3	1	2.826	.094	.013
	Factor4	1	3.139	.078	.014
Ethnicity	Factor1	9	2.652**	.006	.098
	Factor2	9	2.451*	.011	.091
	Factor3	9	1.457	.165	.056
	Factor4	9	3.585***	.000	.128
Current Study Year	Factor1	4	1.701	.151	.030
	Factor2	4	.074	.990	.001
	Factor3	4	.143	.966	.003
	Factor4	4	.473	.756	.009
Major	Factor1	18	.708	.801	.055
	Factor2	18	1.380	.143	.101
	Factor3	18	1.772*	.030	.127
	Factor4	18	.899	.580	.069
Hours of Using	Factor1	4	2.872*	.024	.050
Computer for	Factor2	4	2.519*	.042	.044
Learning per Week	Factor3	4	4.535**	.002	.076
	Factor4	4	2.068	.086	.036
Years of Playing	Factor1	3	.311	.817	.004
Online/Video/Mobile	Factor2	3	.216	.885	.003
Games	Factor3	3	1.131	.337	.015

	Factor4	3	.774	.510	.010
Hours of Playing	Factor1	4	1.098	.358	.020
Online/Video/Mobile	Factor2	4	1.146	.336	.020
Games	Factor3	4	.820	.514	.015
	Factor4	4	3.227*	.013	.055
Gamer Type	Factor1	6	.367	.899	.010
	Factor2	6	.581	.745	.016
	Factor3	6	2.970**	.008	.075
	Factor4	6	.857	.527	.023
Error	Factor1	220			
	Factor2	220			
	Factor3	220			
	Factor4	220			
Total	Factor1	273			
	Factor2	273			
	Factor3	273			
	Factor4	273			
Corrected Total	Factor1	272			
	Factor2	272			
	Factor3	272			
	Factor4	272			

Note. The significance level is .05. Main effect is statistically significant at $p < .05$. * $p < .05$, ** $p < .01$, *** $p < .001$.

Appendix J

Pairwise Comparisons among Different Ethnical Groups on Each Factor in the Modified Survey

Dependent Variable	(I) Ethnicity	(J) Ethnicity	ΔM (I-J)	Sig.
Factor1	White	Chinese	0.28	.000*
		The Others	0.08	1.000
	Chinese	White	-0.28	.000
		The Others	-0.20	.116
	The Others	White	-0.08	1.000
		Chinese	0.20	.116
Factor2	White	Chinese	0.21	.019
		The Others	-0.10	1.000
	Chinese	White	-0.21	.019
		The Others	-0.32	.068
	The Others	White	0.10	1.000
		Chinese	0.32	.068
Factor3	White	Chinese	0.05	1.000
		The Others	-0.04	1.000
	Chinese	White	-0.05	1.000
		The Others	-0.10	1.000
	The Others	White	0.04	1.000
		Chinese	0.10	1.000
Factor4	White	Chinese	0.37	.000*
		The Others	0.02	1.000
	Chinese	White	-0.37	.000
		The Others	-0.35	.003
	The Others	White	-0.02	1.000
		Chinese	0.35	.003*

Note. Based on observed means. ΔM = mean difference between I and J. Bonferroni correction was used to control Type I error. The new adjusted significance level is .017. *. The mean difference is significant at the .017 level.

Appendix K

Finalized Instrument – Pleasurable Learning Experiences Scale

Pleasurable Learning Experiences Scale

Instruction: *Please indicate your agreement with each fun and playful learning experience by choosing from 1 – Strongly Disagree, 2 – Disagree, 3 – Agree, and 4 – Strongly Agree.*

Preferences for Instructions:

1. I like being able to check my progress.
2. I like instructors to provide clear rubrics.
3. I like instructors to provide complete syllabus.
4. I like instructors to give us guidance/directions when we are doing in-class activities.
5. I like instructors to share class materials.
6. I like learning those knowledge and skills that are directly related to the final exam of the course.
7. I like receiving instructors' comments on my assignments.
8. I like getting quick feedback on my performances from the instructors.

Preferences for Instructors' Teaching Styles:

1. I like instructors to make jokes and tell stories during instruction.
2. I like instructors to use clever pictures and relevant quotations during instruction.
3. I like instructors to use videos during instructions.

Attitudes towards Activities:

1. I like interactive in-class activities.
2. I like instructors to use interactive technology (e.g., i-clickers and social media tools).
3. I like collaborating with my classmates.
4. I like building personal connections with my classmates.
5. I like trying new things.
6. I like challenges.

Preferences for Learning Effectiveness:

1. I like building personal connections with instructors.
2. I like instructors' attentions in class.
3. I like reviewing content covered in class.
4. I am self-motivated.
5. I like instructors being accessible.

Appendix L

Demographic Questions Used in the Finalized Instrument

Q1 What is your age?

- ❍ 18-19
- ❍ 20-21
- ❍ 22-23
- ❍ 24+

Q2 What is your gender?

- ❍ Male
- ❍ Female

Q3 Are you of Hispanic, Latino, or Spanish origin?

- ❍ No, not of Hispanic, Latino, or Spanish origin
- ❍ Yes, Mexican, Mexican Am., Chicano
- ❍ Yes, Puerto Rican
- ❍ Yes, Cuban
- ❍ Yes, another Hispanic, Latino, or Spanish origin -- Print origin _____

Q4 What is your Ethnicity? (Mark one or more boxes)

- ❑ White
- ❑ Black or African American
- ❑ American Indian or Alaska Native -- Print name of enrolled or principal tribe (4)
  _____
- ❑ Asian Indian
- ❑ Chinese
- ❑ Filipino
- ❑ Japanese
- ❑ Korean
- ❑ Vietnamese
- ❑ Native Hawaiian
- ❑ Guamanian or Chamorro
- ❑ Samoan
- ❑ Other Asian -- Print race _____
- ❑ Other Pacific Islander -- Print race _____
- ❑ Some other race -- Print race _____

Q5 Are you an international student in the United States?

- ❍ Yes
- ❍ No

Q6 What is your current year of study?

- ❍ 1st year
- ❍ 2nd year
- ❍ 3rd year
- ❍ 4th year
- ❍ 5th + year

Q7 What is your major? (Mark one or more boxes)

- ❑ Architecture
- ❑ Arts
- ❑ Business
- ❑ Education
- ❑ Engineering
- ❑ Health Professions
- ❑ Health, Sport, and Exercise Sciences
- ❑ Humanities & International Studies
- ❑ Journalism & Mass Communications
- ❑ Law (Pre-Law)
- ❑ Medicine (Pre-Med)
- ❑ Music
- ❑ Natural Science & Math
- ❑ Nursing
- ❑ Pharmacy
- ❑ Public Affairs & Administration
- ❑ Social & Behavioral Sciences
- ❑ Social Welfare
- ❑ Undecided

Q8 How long do you use computers for learning?

- ❍ Never
- ❍ 1-3 hours per week
- ❍ 4-6 hours per week
- ❍ 7-9 hours per week
- ❍ 10 + hours per week

Q9 Please check the experiences you have had before (you can choose multiple items):

- ❑ Massive multiplayer online games
- ❑ Massive open online courses (MOOCs)
- ❑ Online shopping
- ❑ Social network sites (e.g., Facebook, Twitter, LinkedIn)
- ❑ None of above

Q10 How long have you played video/online/mobile games?

- ❍ 1-3 years
- ❍ 4-6 years
- ❍ 7+ years
- ❍ Never

Q11 How long do you play video/online/mobile games per week?

- ❍ 1-5 hours
- ❍ 6-10 hours
- ❍ 11-15 hours
- ❍ 16 + hours
- ❍ Never

Q12 Which gamer type do you think best describe you?

- ❍ Story-driven solo gamer (who play alone and are motivated by well-written story-based games)
- ❍ Social gamers (who view their gaming as a social activity)
- ❍ Solo-limited gamers (who casually play a wide range of single-player games)
- ❍ Hardcore online gamers (who do not put restraints upon their gaming and hate being interrupted)
- ❍ Control/identity solo gamers (who particularly like playing story-driven games that allow for elements of character choice or development)
- ❍ Casual gamers (who play graphically good, short games or mission-based games when they have time to do so)
- ❍ None of above